Computed Tomography Based Lung Function Mapping Using Deep Learning for Functional Lung Avoidance Radiation Therapy

REN, Ge

PhD

The Hong Kong Polytechnic University

2021

The Hong Kong Polytechnic University

Department of Health Technology and Informatics

Computed Tomography Based Lung Function Mapping Using Deep Learning for Functional Lung Avoidance Radiation Therapy

REN, Ge

A thesis submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

January 2021

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

_____REN, Ge_____(Name of Student)

Abstract

Background: Functional lung avoidance radiation therapy (FLART) is an emerging technique that selectively minimizes dose delivery to the high functional lung while favoring dose deposition in the low functional lung based on the information from pulmonary function imaging, which measures the air breathing (ventilation) and blood circulation (perfusion) within the lung. However, the currently established lung function imaging methods require additional scan(s) that are typically not prescribed for radiotherapy, and can be resource-demanding, inconvenient, costly, and technically challenging for radiotherapy treatment planning.

Purpose: This study aims to develop a deep learning-based computed tomography (CT) function mapping (CTFM) method that is able to synthesize lung perfusion images and lung ventilation images from the CT domain for FLART.

Methods and Materials: In the first part, we developed a deep learning-assisted framework to extract features from 3D CT images and synthesize perfusion and ventilation as estimations of lung function using CT images. This CTFM framework includes three parts: image preparation, image processing pipeline, and our proposed convolutional neural network (CNN). Image preparation consists of a series of morphological operations to decrease computational consumption and remove noise. Image processing aims to enhance the feature robustness of CT images and standardize SPECT perfusion images to be suitable labels for CNN mapping. It consists of SPECT normalization, SPECT discretization, CT contrast enhancement, and CT defect

enhancement. A three-dimensional (3D) attention residual neural network (ARNN) was constructed to extract textural features from the CT images and reconstruct corresponding functional images. Three components (residual module, ROI attention, and skip attention) were used to improve the performance of the neural network. In the second part, we investigated the effects of each framework component and setting for CT-to-perfusion translation. The first cohort of 42 pulmonary macro-aggregated albumin SPECT/CT perfusion scans were retrospectively collected from the hospital. Ablation experiments and comparison experiments were performed by changing each element of the framework and analyzing the testing performances. In the third part, a total number of 73 pulmonary MAA SPECT/CT images were collected for performance evaluation. A quantitative comparative analysis between the predicted perfusion and SPECT perfusion was conducted to evaluate the overall performance. To assess the function-wise concordance, the Dice similarity coefficient (DSC) was computed to determine the similarity of the low/high functional lung volumes. In the fourth part, a total number of 13 Technegas SPECT/CT ventilation scans were collected from patients for suspected lung disease in the hospital. This dataset was used to evaluate the feasibility of CT-to-ventilation using the proposed CTFM framework. The analysis metric mainly includes the Spearman's correlation coefficient (R) and structural similarity index measure (SSIM), accounting for statistical and perceptual image similarity, respectively.

Results: Major results of our investigation experiments of CTFM framework showed that the removal of the CT contrast enhancement component in the image processing part resulted in the largest drop in framework performance compared to the optimal

performance (~11%). In the CNN component part, all the three components (residual module, ROI attention, and skip attention) were approximately equally important to the framework performance; removing either one of them resulted in a 3-5% decline in performance. In the CNN configuration part, a combination adoption of both Adam optimizer and loss functions of binary cross-entropy (BCE) and Spearman's correlation coefficient led to the best framework performance. Our proposed framework achieved ~4% higher overall performance and 4.5-fold higher computation efficiency, compared to the U-Net model. The evaluation of the voxel-wise agreement showed a moderate-to-high voxel value correlation (0.6733 ± 0.1728) and high structural similarity (0.7635 ± 0.0697) between the SPECT and DL-CTFM predicted perfusions. The evaluation of the function-wise concordance obtained an average DSC value of 0.8183 ± 0.0752 for high-functional lungs, ranging from 0.5819 to 0.9255, and 0.6501 ± 0.1061 for low-functional lungs, ranging from 0.2405 to 0.8212. Ninety-four percent of the test cases demonstrated high concordance (DSC > 0.7) between the high functional volumes contoured from the predicted and ground-truth perfusions. For the preliminary result in ventilation feasibility study, ventilation images generated using our CTFM framework in the testing group had an average R of 0.6336 ± 0.0815, average SSIM of 0.7391 ± 0.0681, average DSC of 0.7089 ± 0.0438 for lung functional lung, and average DSC of 0.7265 ± 0.0473 for high functional lung.

Conclusion: In this study, we developed a novel DL-CTFM framework for estimating lung functional images from the CT domain using a 3D ARNN. The deep convolutional neural network, in conjunction with image processing pipeline for feature enhancement, is capable of feature extraction from CT image for perfusion synthesis. This CTFM

framework provides insights for relevant research studies in the future and enables other researchers to leverage for the development of optimized CNN models for CTFM. In the CT-to-perfusion/ventilation translation, the CTFM framework yields moderate-to-high voxel-wise approximations of lung function. To further contextualize these results toward future clinical application, a multi-institutional large-cohort study is warranted.

Publications

Journal Articles

1. **G. Ren**, W. Ho, C. Yin, H. Xiao, T. Li, J. Qian and J. Cai, Deep Learning Based Computed Tomography Perfusion Mapping (DL-CTPM) for Functional Lung Avoidance Radiation Therapy (FLART), *International Journal of Radiation Oncology, Biology, Physics (IF: 5.859).* (**Accepted**)

2. **G. Ren**, S. Lam, Z. Jiang, H. Xiao, A. Cheung, W. Ho, J. Qin and J. Cai, Investigation of a Novel Deep Learning-Based Computed Tomography Perfusion Mapping Framework for Functional Lung Avoidance Radiotherapy. *Frontiers Oncology (IF:4.848).* (**Accepted**)

3. **G. Ren**, H. Xiao, S. Lam, D. Yang, T. Li, X. Teng, J. Qin and J. Cai, CT-based Bone Suppression in Chest Radiographs Using Cascade Convolutional Neural Network: A Feasibility Study. *Quantitative Imaging in Medicine and Surgery (IF:3.226).* (**Accepted**)

4. **G. Ren**, S. Lam, R. Ni, D. Yang, J. Qin, and J. Cai, The Effectiveness of Data Augmentation for Bone Suppression Chest Radiograph using Convolutional Neural Network. (**Paper in editing**)

5. **G. Ren**, W. Li, H. Xiao, J. Cai, Roadmap paper on Artificial Intelligence in Radiation Oncology, Chapter of Image Synthesis. (**Paper in editing**)

6. **<u>G. Ren</u>**, W. Ho, J. Qian, J. Cai, Deriving Lung Perfusion Directly from CT Image Using Deep Convolutional Neural Network: A Preliminary Study, Springer International Publishing. D. Nguyen, L. Xing, S. Jiang (Eds.), *Artificial Intelligence in Radiation Therapy* (pp. 102-109), New York, United State: Springer International Publishing, (2019).

7. T. Peng, Y. Wu, C. Wang, J. Zhao, C. Tang, Z. Ye, **<u>G. Ren</u>**, J. Cai, H-Pro, Seg: Hybrid Prostate Segmentation Algorithm in Transrectal Ultrasound Images using Redefined Closed K-Segment Principal Curve and Improved Differential Evolution-based Artificial Neural Network, *IEEE Transactions on Pattern Analysis and Machine Intelligence (IF: 17.861)*. (**Under review**)

8. T. Li, D. Cui, **<u>G. Ren</u>**, E. Hui and J. Cai, Investigation of the Effect of Acquisition Schemes on Time-Resolved Magnetic Resonance Fingerprinting, *Physics in Medicine and Biology (IF:3.030),* (**Under review**).

9. X. Tang, Y. Shen, Y. Meng, L. Hou, C. Zhou, C. Yu, H. Jia, W. Wang, **<u>G. Ren</u>**, J. Cai, X. Li, H. Yang, F. Kong, Radiation-induced lung damage in patients treated with stereotactic body radiotherapy after EGFR-TKIs: Is there any difference from stereotactic body radiotherapy alone, *Annals of Palliative Medicine, (IF:2.477)*. (**Accepted)**

10. H. Xiao, **<u>G. Ren</u>**, J. Cai, A review on 3D deformable image registration and its application in dose warping. *Radiation Medicine and Protection (IF: N/A)*, ISSN 2666-5557, (2020).

Conference Abstract

1. **G. Ren**, W. Ho, H. Xiao, A. Cheung, J. Qin, J. Cai, Evaluation of Deep Learning Based CT Perfusion Imaging (DLCT-PI) for Functional Lung Avoidance Radiation Therapy (FLART). The 23rd Annual Meeting of Chinese Society of Clinical Oncology (CSCO), 2020, Beijing, China. **(Top 4 of 60 Submissions, Oral Presentation)**

2. **G. Ren**, W. Ho, H. Xiao, A. Cheung, J. Qin, J. Cai, CT-based Lung Perfusion Mapping Using Attention Residual Neural Network for Functional Avoidance Radiation Therapy. Final List of J.R. Cunningham Young Investigator Symposium in Joint AAPM | COMP, 2020. **(Top 10 of 396 Submissions, Oral Presentation)**

3. **G. Ren**, W. Ho, J. Qin, J. Cai, Deriving Lung Perfusion Directly from CT Image Using Deep Convolutional Neural Network: A Preliminary Study. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2019. Shenzhen, China. **(Oral Presentation)**.

4. **G. Ren**, Y. Zhang, F. Yin, L. Ren. Enhancing on-board image contrast using prior images and deformable registration for target localization: a feasibility study, 2019 AAPM Annual meeting, Nashville, TN, United State. (**Abstract**)

5. **G. Ren**, J. Cai. Assessing the Feasibility of Using Deformable Image Registration for On-board Multi-Modality Based Target Localization in Radiation Therapy, AAPM Seminar, 2018, Shanghai, China. (**Abstract**)

6. **<u>G. Ren</u>**, J. Qin, W. Ho, J. Cai.基于深度卷积神经网络的肺功能图像合成. The 16th National Congress of Radiation Oncology (NSRO), 2019, Shenzhen, China. (**Abstract**)

7. **<u>G. Ren</u>**, J. Qin, W. Ho, J. Cai.基于深度卷积神经网络的肺通气显像. Chinese Society of Medical Physics (CSMP), 2019, Zhengzhou, China. (**Abstract**)

Acknowledgement

First and foremost, I would like to express my sincere gratitude and appreciation to my chief supervisor, Prof. Jing Cai, and co-supervisor, Dr. Jing Qin, for providing instructive guidance and spiritual encouragement throughout my research. Their dynamism, vision, and motivation have deeply inspired me. Besides, I would like to express my special thanks to Dr. Jing Cai. He has taught me the methodology to carry out the research and the skills to present the work. It was a great honor to work and study under his guidance.

I am extending my special thanks to Dr. Waiyin Ho, for his strong support in this project, and his valuable suggestions during this research.

I am very thankful to Dr. Xiangyan Fiona Chen, Dr. Michael T.C. Ying, Dr. Jung Sun Yoo, Dr. Helen K.W. Law, and Dr. Shea-ping Yip, for their professional advice and help.

I would like to express my thanks to my group members for their generous help and support in my research. Special thanks to Mr. Tian Li, Mr. Haonan Xiao, Mr. Saikit Lam, Mr. Jiang Zhang, Mr. Xinzhi Teng, Ms. Tingting Yu, for their valuable advice in starting the research.

I would like to extend my thanks to my friends and all the other people who have supported me directly or indirectly for giving me suggestions, encouragement, and help. Special thanks to HTI staffs for their constant and prompt assistances.

Finally, I would not forget to thank my parents and sister for their unconditional love and sacrifices for my education. Special thanks to my girlfriend for her love, understanding, prayers, and continuing support during my research. Their encouragement and understanding are always my greatest motivation.

Contents

XIV

List of Tables

List of Figures

Abbreviations

2D	Two dimensional
3D	Three dimensional
AdaGrad	Adaptive gradient algorithm
ADAM	Adaptive moment estimation optimizer
AI	Artificial intelligence
ANN	Artificial neural network
ARNN	Attention residual neural network
CBCT	Cone-beam computed tomography
chf	Cumulative histogram function
CNN	Convolutional neural network
ConT	Transposed convolution
COPD	Chronic obstructive pulmonary disease

COVID	Coronavirus disease
CRT	Concurrent chemotherapy
CT	Computed tomography
CTFM	CT Function Mapping
CTFM	CT function mapping
CTVI	Computed tomography ventilation imaging
CTVM	CT ventilation mapping
CXR	Chest x-ray
DCNN	Deep convolutional neural network
DIR	Deformable image registration
DL	Deep learning
DL-CTFM	Deep learning-based CT function mapping
DNA	Deoxyribonucleic acid

DSC	Dice similarity coefficient
DTPA	Diethylene triamine peracetic acid
FCN	Fully convolutional network
FLART	Functional lung avoidance radiotherapy
Ga	Gallium
GAN	Generative adversarial network
HNTH	Henan Tumor Hospital
HP	Hyperpolarized
HU	Hounsfield units
IRB	Institutional Review Boards
IV	Injection of the vein
Kr	Krypton
MAA	Macro aggregated albumin

MLD	Mean lung dose
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MSE	Mean square error
NSCLC	Non-small cell lung cancer
NTCP	normal tissue complication probability
PE	Pulmonary embolism
PET	Positron emission tomography
PReLU	Parametric rectified linear unit
QMH	Queen Mary Hospital
R	Spearman correlation coefficient
ReLU	Rectified linear unit
ResConv	Residual convolution

ROI	Region of interest
ROI Att	Region of interest attention module
RP	Radiation pneumonitis
RT	Radiation therapy
RTOG	Radiation Therapy Oncology Group
SCLC	Small cell lung cancer
SD	Standard deviation
SGD	Stochastic gradient descent
Sigmoid	Sigmoid function
SkipAtt	Skip attentions
SPECT	Single-photon emission computed tomography
SSIM	Structural similarity index measure
StriConv	Strided convolution

Tc	Technetium
V/Q	Ventilation/Perfusion
V20	Volume receiving more than 20 Gy
VAMPIRE	Ventilation and medical pulmonary image registration evaluation
WHO	World Health Organization
Xe	Xenon

1.   Literature Review

1.1   *Radiation Therapy for Lung Cancer*

1.1.1   Brief Introduction of Lung Cancer

The two lungs are a pair of half cone-shaped organs of the respiratory system located on each thoracic cavity side. Air enters the lungs via the branches of the trachea. In the microscopic air sacs end of the lung, oxygen from the inhaled air is absorbed into the bloodstream [1]. Lung cancer causes the uncontrolled growth of cells in lung tissues. The lungs exchange up to 7200 L of air per day, which may contain pathogens and other potentially harmful elements. These exposures affect the cell's normal functions, including cell proliferation, programmed cell death, and deoxyribonucleic acid (DNA) repair. As more damage accumulates, the risk for cancer increases. Eventually, lung cancer may develop after genetic damage to DNA and epigenetic changes [2]. Because of smoking and air pollution, the lung cancer incidence rate increased in the last decades [3]. According to the report of the World Health Organization (WHO), lung cancer is the most commonly occurring cancer in adults, and is also the most common cause of cancer-related death [4]. In estimation, there were 2 million new cases in 2019 [5]. In 2018, there were 1.7 million death cases, accounting for 11.6% of the total cancer patients worldwide. The number of patients who died from lung cancer equals the add up of patients who died of the prostate, breast, and liver cancers in 2018 (Figure 1). In China, lung cancer is also the most commonly diagnosed cancer type. There are 7774,323 new cases in 2018, occupying 18.1% of the total cancer patients [6].
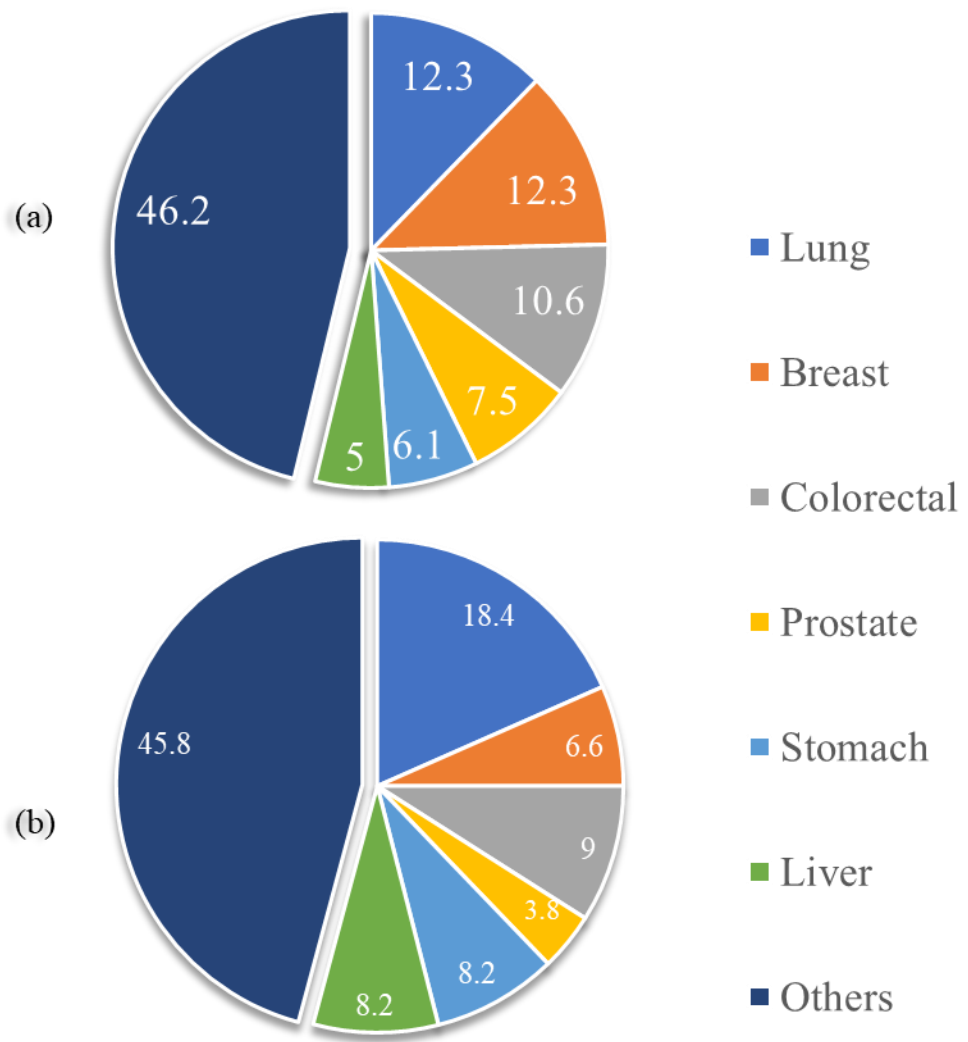
Figure 1 Global cancer incidence (a) and mortality (b) in 2018 [5].

The diagnostic tests for lung cancer include chest X-ray (CXR), high-resolution lung CT, and tissue biopsy. The CXR may reveal an apparent mass. Lung cancer appears as a solitary nodule on the CXR. The chest CT scan may reveal a spiculated mass, which is highly suggestive of lung cancer and is also used to provide more information about the disease's type and extent. However, CT imaging cannot be used frequently since the extended radiation exposes an increased risk of lung cancer. Lung biopsy is an examination of tissue removed from the lung to discover the presence, cause, or extent of lung cancer [7].

Lung cancer treatment depends on tumor type and its stage. Over 90% of lung cancer are small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLC, also named oat cell cancer, takes up approximately 10% to 15% of diagnosed lung cancer. SCLC grows faster than NSCLC. About 70% of SCLC patients have spread cancer at the time of diagnosis. Since it proliferates quickly, it tends to respond well to chemotherapy and radiotherapy. Approximately 80 to 85% of lung cancer patients were diagnosed with NSCLC, the main subtypes of which are large cell carcinoma, squamous cell carcinoma, and adenocarcinoma (Figure 2). The treatment of these subgroups is similar because they have a similar prognosis. Compared with small cell carcinoma, NSCLC is relatively unresponsive to chemotherapy. The early-stage NSCLC patients require comprehensive treatment, and the primary treatment is surgical resection followed by radiation therapy (RT) [8]. About 30% to 50% of NSCLC patients present with locally advanced NSCLC (stage III). The current standard treatment for locally advanced NSCLC is concurrent chemotherapy (CRT) [9, 10]. Radiotherapy in consort with surgery and chemotherapy plays a crucial role in the treatment of lung cancer [11].
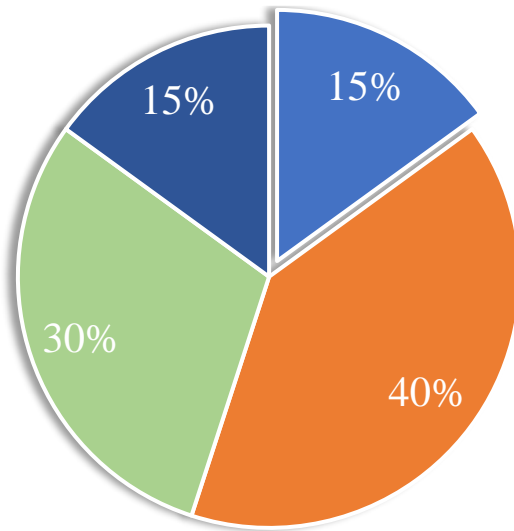
Figure 2 The main types of lung cancer.

1.1.2   Radiation-Induced Lung Injury

Lung cancer radiation therapy uses a high-energy photon beam to break the double-strand structure in DNA and kill the cancer cells. However, excess dose to healthy lung tissue causes lung injury complications, which is common treatment-related toxicity. This injury manifests in a biphasic manner: radiation pneumonitis (RP) in the acute term and pulmonary fibrosis in the long term. RP is a kind of pulmonary inflammation caused by the X-ray irradiation to the lung (Figure 3). When high radiation does are used, or large lung areas are exposed to radiation, RP frequently develops in 1 to 3 months after radiation treatment. Chronic RP can lead to permeant scarring of the lung, referred to as pulmonary fibrosis (PF). PF is a disease that occurs when lung tissue becomes damaged and scarred. The symptoms of lung damage can show months or even years after the initial treatment. The fibrosis tissues become thickened and stiff, making it more difficult for the lung to work properly. As the disease worsens, it progressively leads to shortness of breath. The damaged lung tissue cannot be repaired, which decreases the patient's life quality [12, 13].

Radiation-induced lung injury is also a major cause of treatment-related morbidity. As high as 40% of patients developed clinically significant pneumonitis after lung radiation therapy [14]. The side effects impair the treatment outcome and decrease the survival rate of the patients [15]. Study shows radiation-induced lung injury decreases the 5-year survival rate from 18-36% to 5-14% for lung cancer patients [16]. In clinical practice, parameters such as mean lung dose (MLD) and lung tissue volume receiving more than 20 Gy (V20) are used to judge whether radiation treatment plans are safe. However, both parameters are imperfect tools for predicting radiation-induced

pulmonary toxicity. Patients with MLD and V20 at levels considered to be safe still have to 20% risk of experiencing worsening shortness of breath [17].

The RP also remains major dose-limiting toxicity and precludes dose escalation for better local control of lung cancer. The modeling study has demonstrated that higher doses of radiotherapy should be correlated with improved tumor control [18]. However, dose-escalation also increases the risk of excess pulmonary toxicity. The clinical trial of Radiation Therapy Oncology Group (RTOG)-0617 compared the overall survival of locally advanced NSCLC treated by standard radiotherapy (60 Gy) and high dose radiotherapy (74 Gy), followed with concurrent chemotherapy. Overall survival at 18-months was 66.9% in the standard radiotherapy and 53.9% in high dose radiotherapy, indicating inferior survival with dose-escalation [19].
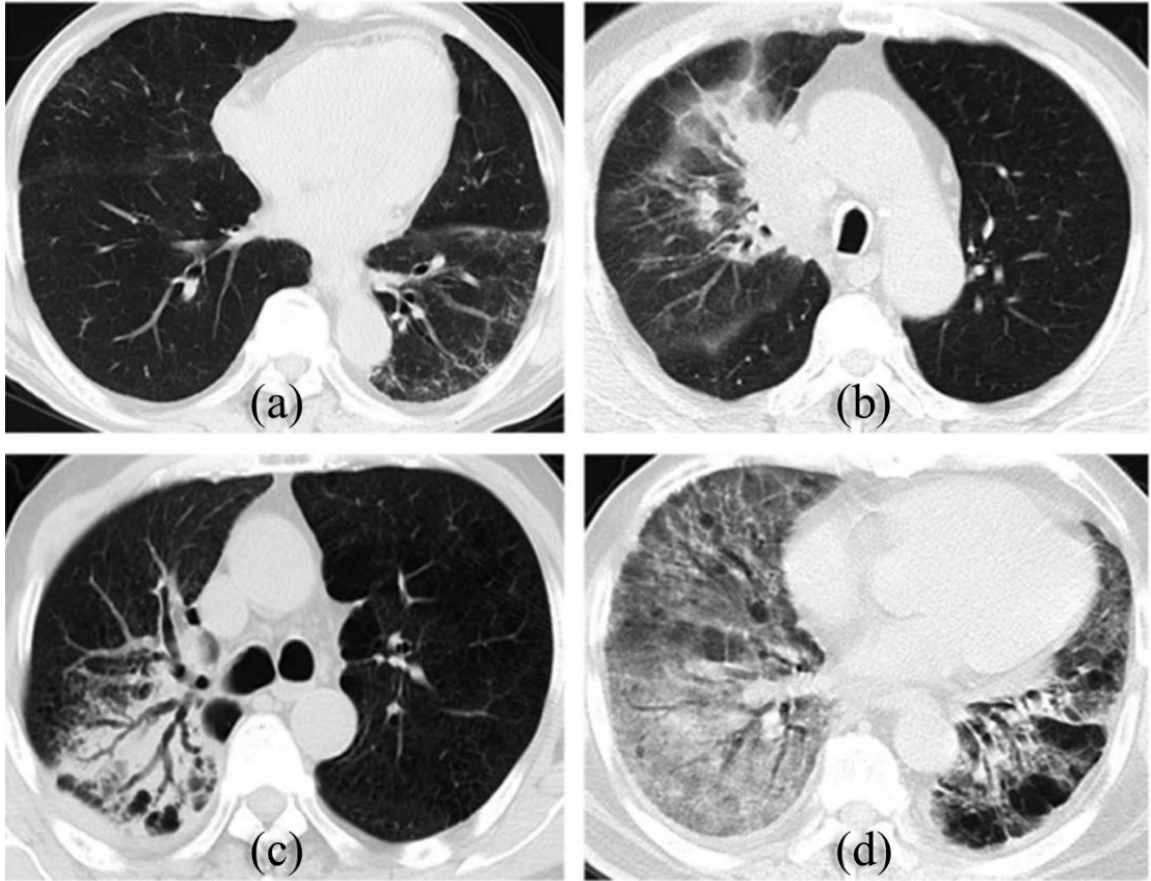
Figure 3 Radiation pneumonitis (RP) in CT images: (a) Grade 1, (b) Grade 2, (c) Grade 3, and (d) Grade 4 [20].

1.1.3   Functional Lung Avoidance Radiation Therapy

Current radiotherapy is based on anatomic information provided by planning CT scan. Planning CT provides a complete 3D view of the patient's anatomy, which indicates the tumor and surrounding organs. The pixel values in CT scans are in Hounsfield units (HU), which can be converted to electron density for dose calculations [21]. As shown in Figure 4, the tumor region and the region immediately surrounding it are planned to receive the maximum radiation dose (100% dose). Radiotherapy treatment planning aims to maximize radiation dose to the tumor while minimizing the dose delivered to the normal lung tissue. However, the extend of radiation-induced injuries depend on the dose and volume [11]. Surrounding normal lung regions receives radiation doses that decrease with increasing distance from the tumor. However, CT is not an ideal tool to guide treatment planning. CT-based treatment planning only allows anatomic correlation and assumes homogeneous radiation dose-response of normal tissues. For the radiation therapy of lung cancer, current treatment planning considers that lung regions have a similar functional response. It is not responsible for the fact that lung tissue can be heterogeneous, especially in lung cancer patients whose lungs are frequently characterized by large variations of regional lung functions. These uniform function cases are rare in lung tumor patients [22]. In addition, lung tumors (especially centrally located tumors) can affect the function of adjacent lung tissue. This heterogeneity can be important in radiation planning because low functional lung tissue is expected to have a different reaction to radiation beams than high functional lung tissue. For lung cancer patients considered radiotherapy, the zone of radiation should be minimized to the tumor volume, and irradiation of normal tissue should be avoided.
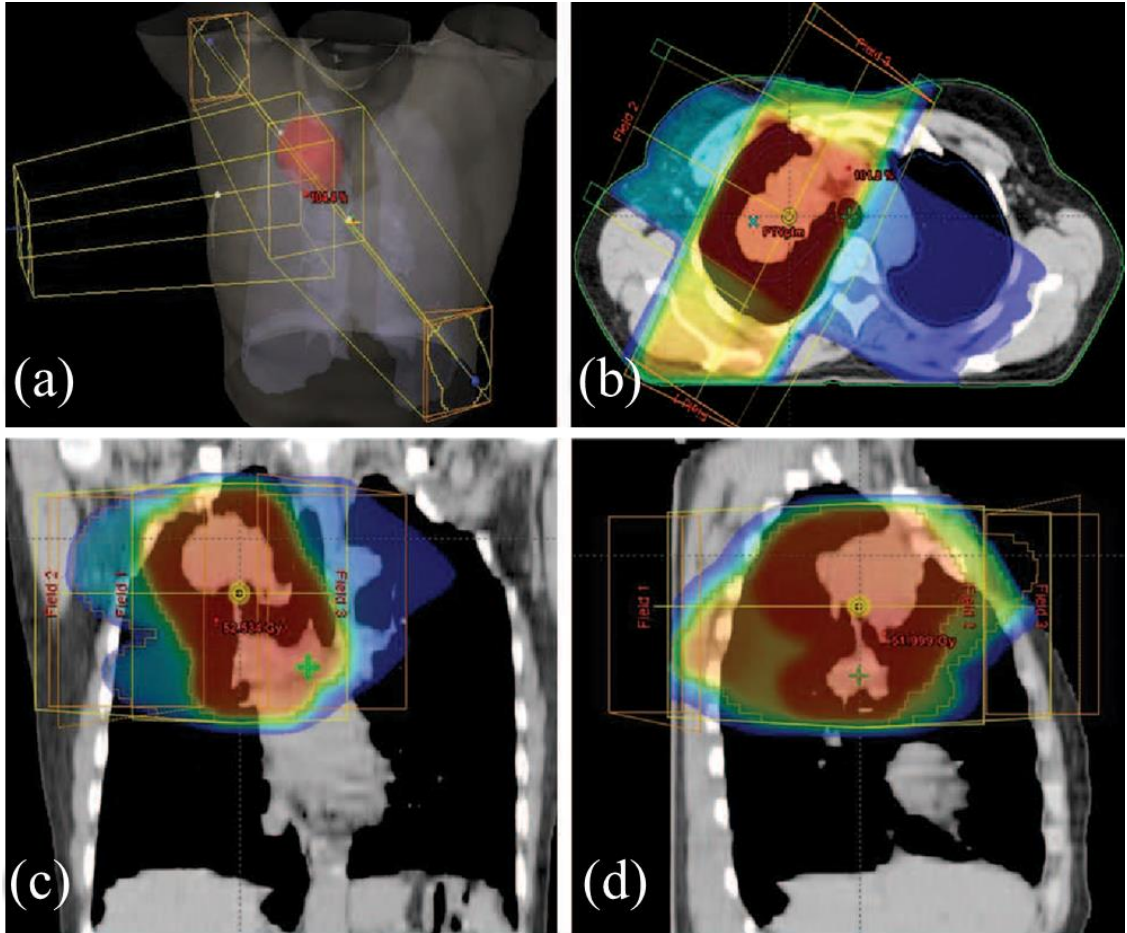
Figure 4 An illustration of radiation therapy planning for lung cancer. (a) three

radiation beams delivered to the tumor region. (b) axial, (c) coronal, and (d) sagittal

CT images show the spatial distribution of isodose curves [23].

Radiation therapy treatment planning should be able to exploit the regional function difference. Functional lung avoidance radiotherapy (FLART) is an emerging technique that selectively avoids irradiating highly-functional lung regions and favoring dose deposition in low functional lung regions [18]. FLART holds great promise to reduce radiation-induced lung injury (Figure 5). Studies have found that functional avoidance radiotherapy reduced the risk of Grade ≥ 2 RP by 5%–8% on average for the lung cancer population (up to 52% for individual patients) compared to standard radiotherapy [11], with a reduction of the mean dose to the functional lung by ~ 2 Gy without significantly increasing the doses to other structures [22].

Currently, there are three ongoing clinical trials in the United States (NCT02528942, NCT02308709, and NCT02843568) investigating the clinical efficacy of lung ventilation images for functionally guided radiotherapy of lung cancer. In addition, Matuszak et al. integrated SPECT perfusion in dose optimization of 15 patients and found that the mean dose in the high functional lung region decreased from 12.6 ± 4.9 Gy to 9.9 ± 4.4 Gy [24]. Faught et al. found that lung ventilation-guided plan optimization reduced to grade 2 and above normal tissue complication probability (NTCP) up to 7.8% [25]. Waxweiler et al. observed an average decrease of the mean dose to the functional lung by 2.8 Gy in CT ventilation imaging (CTVI) guided planning [26]. Yamamoto et al. reported a 5.0% decrease in the functional lung dose by CTVI-guided planning [27]. Functionally guided radiotherapy was found to improve dose-volumetric outcomes for functional lung. As high as 40% of patients developed clinically significant pneumonitis after lung radiation therapy. The pneumonitis also remains major dose-limiting toxicity and precludes dose escalation for better local control. The side effects impair the treatment outcome and decrease the survival rate of

the patients [15]. Radiation pneumonitis decreases the 5-year survival rate from 18-36% to 5-14% [16]. Current radiotherapy is based on anatomic imaging and assumes a homogeneous radiation dose-response of normal tissues. It is not responsible for the fact that lung tissue can be heterogeneous, especially in lung cancer patients whose lungs are frequently characterized by large variations of regional lung functions.
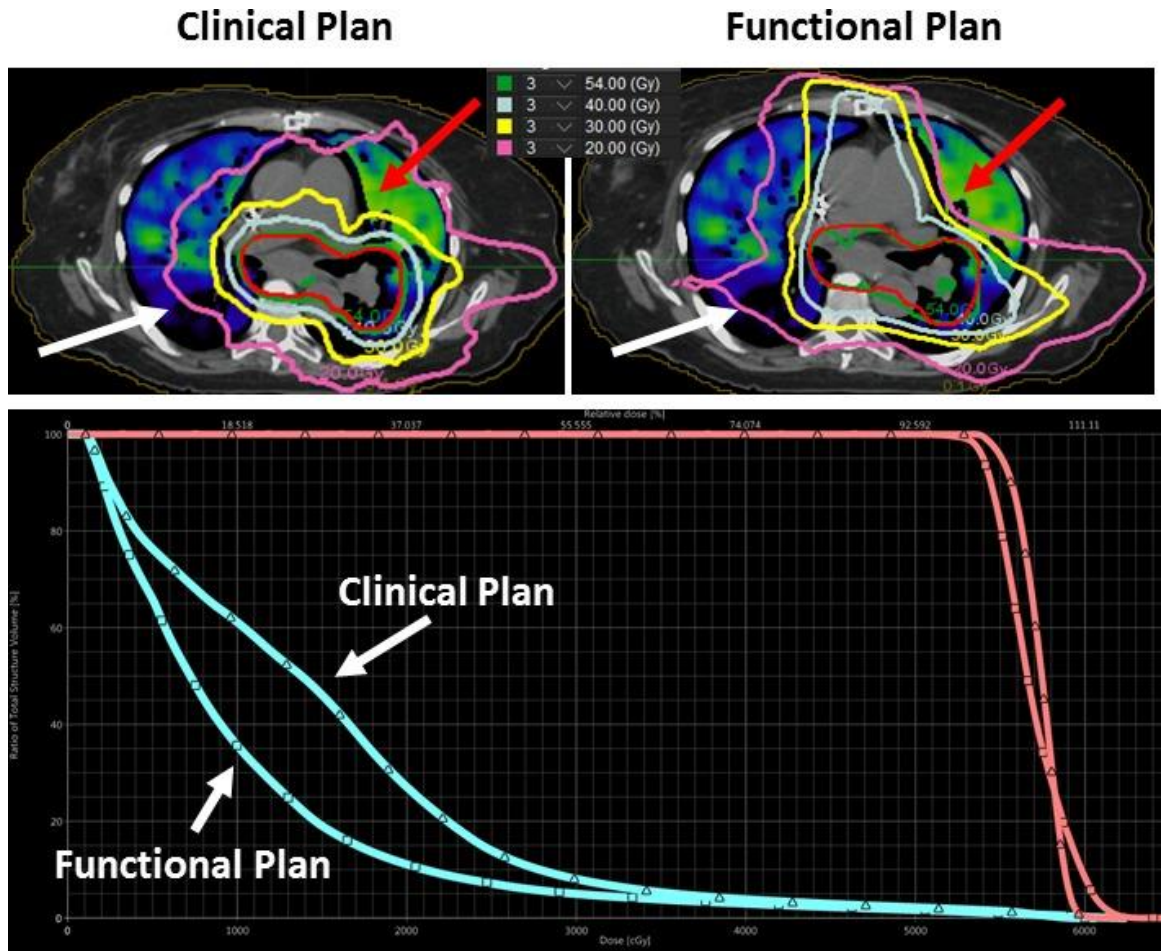
Figure 5 Comparison of treatment planning between a conventional clinical plan and the functional avoidance plan. Each image in the first row displays the planning CT, overlaid by pulmonary functional image, isodose lines, and the DVH curves. The red arrows indicate where the spared functional lung. The white arrows demonstrate how the functional plan deposited higher doses in the low functional region [26].

1.2 Lung Function Imaging

1.2.1 Pulmonary Ventilation/Perfusion (V/Q)

The primary function of the lungs is the gas exchange between the circulatory system and the external environment [28]. In respiration, oxygen from incoming air enters the blood, and carbon dioxide, waste gas from the metabolism, leaves the blood. The ability to exchange gases can be measure by pulmonary ventilation/perfusion (V/Q) imaging, which can reveal the regional distribution or spatial heterogeneity of lung function (Figure 6). A V/Q scan uses radioactive material (radiopharmaceutical) to examine airflow (ventilation) and blood flow (perfusion) in the lungs. The ventilation scan measures how well the air is able to flow through the lungs. The perfusion scan shows where blood flows in the lung.

It is noteworthy that there is a regional vitiation in the V/Q ratio within the lung [28]. From the base of the lung to the apex of the lung, both ventilation and perfusion increase going. The impact of gravity and the weight of the lung expresses itself as an increased pleural pressure at the base lung, reducing the alveolar volume. As a result, the smaller alveolar volume at the base lung is more distensible and more capable of oxygen exchange. For perfusion, gravity increases the blood pressure at the base lung and thus increase the perfusion basal lung area. Besides, perfusion rises to a greater extent than ventilation, lowing the V/Q ratio at the base lung. In general, perfusion increases more than ventilation at the bottom lung, decreasing the V/Q ratios in the bottom lung compared to the apex [29].

In the normal V/Q scans, both lungs should have uniformly distributed radioisotopes. However, a number of conditions can cause abnormal perfusion and

ventilation. A decreased uptake of radioisotope during a perfusion scan indicates a problem with blood flow, including occlusion of the pulmonary arteries. A localized decrease in perfusion scan uptake may indicate pulmonary embolus. Larger areas of decreased perfusion scan uptake may signify a condition such as pneumonitis. A decreased uptake of radioisotope during a ventilation scan may indicate reduced breathing ability or airway obstruction. A decreased ventilation uptake (plus x-ray evidence of consolidation) may indicate pneumonia. Larger areas of poor uptake may indicate damage from chronic smoking or chronic obstructive pulmonary disease (COPD) [30, 31].

Physiologic shunts may also occur if perfusion reaches areas of the lungs that are not ventilated, which may cause by airway obstructions, pulmonary edema, and pneumonia. Asthma is regarded as a "false shunt" because the constriction of the airways in the lungs decreases ventilation due to the tightening of surrounding smooth muscle, resulting in a low V/Q ratio. Pulmonary embolism (PE) is caused by the blockage of the lung artery, resulting in decreased perfusion in alveoli and a high V/Q ratio. The impaired gas exchange may cause hypoxemia in cases of PE [32]. A variety of conditions can affect perfusion. In pulmonary fibrosis, the fibrosis of parenchyma increases the alveolar wall thickness, decreasing the diffusion ability of the lung [33]. Another significant disease that affects perfusion is lung edema. Fluid in the lungs increases the thickness of the alveolar wall and decreases the effective gas exchange volume. The pulmonary edema prevents air from reaching pulmonary capillaries, resulting in perfusion without ventilation and dead alveolar space.
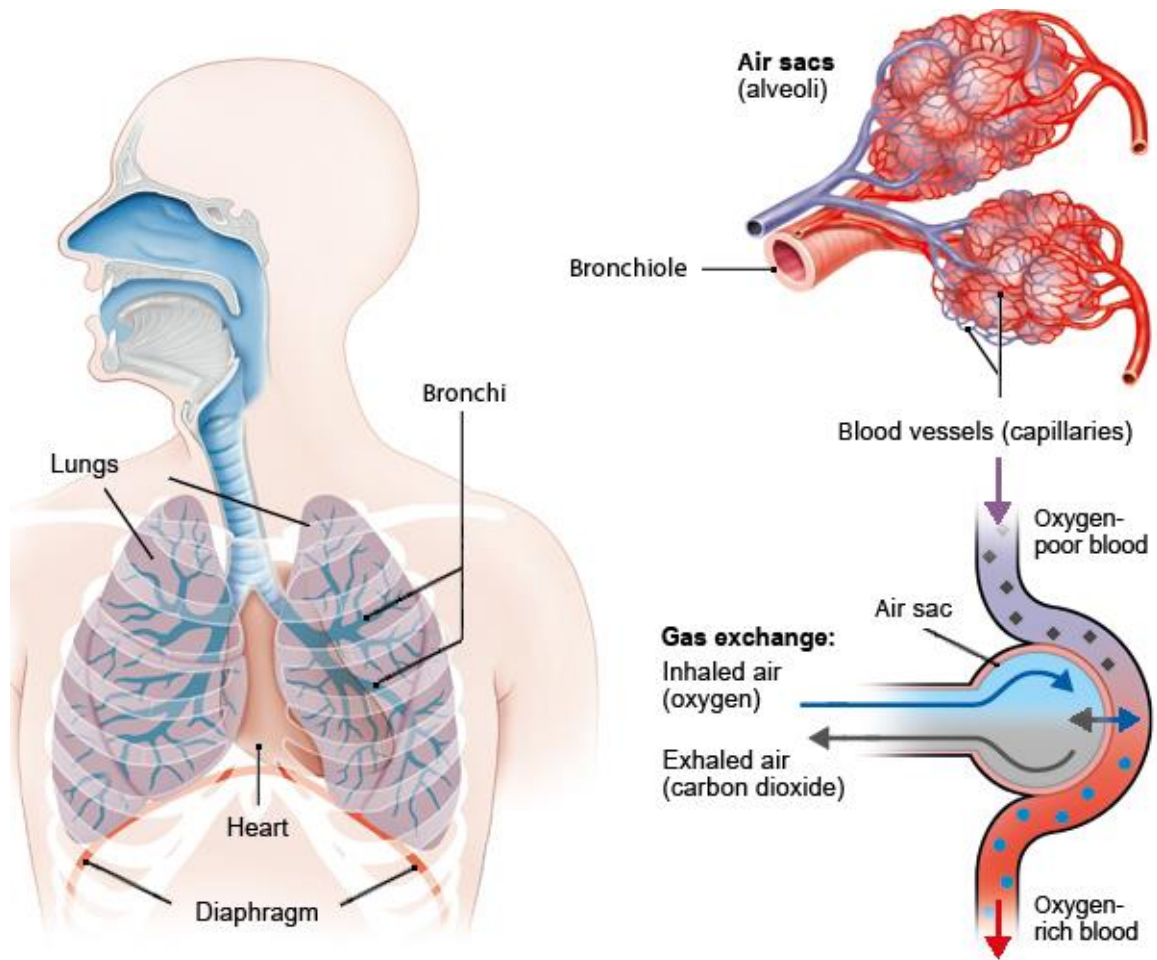
Figure 6 Illustration of the circulation of air (ventilation) and blood (perfusion) in gas exchange [34].

1.2.2 SPECT Ventilation/Perfusion Imaging

Single-photon emission computed tomography (SPECT) ventilation/perfusion (V/Q) scan is the most commonly used lung functional imaging method. It uses scintigraphy and medical isotopes to evaluate the regional circulation of air and blood within the lungs. In clinical practice, SPECT/CT V/Q scanning is a commonly used diagnostic imaging test of pulmonary function [35]. The two scans involve the use of a low-risk radioactive substance that can be traced by a particular type of scanner. The substance will gather at areas of abnormal blood or airflow, which may indicate a blockage in the lung (Figure 7).

In the ventilation scan, the administered radiopharmaceutical ventilation agent is inhaled by the patient. The SPECT scanner collects the signal delivered by the radioisotopes. The inhaled radiopharmaceuticals include three categories: (i) inert gases such as xenon (Xe)-133, krypton (Kr)-81m. (ii) Radio aerosols, such as technetium (Tc)-99m diethylene triamine peracetic acid (DTPA). (iii) Technegas, which is a carbon-coated ^{99m}Tc nanoparticulate aerosol [36-38]. Studies using SPECT perfusion with long-term follow-up provided rationale that regional perfusion reduction is correlated with delivered radiation dose [39].

The perfusion agent includes Tc-99m labeled macro aggregated albumin (MAA), which is also the most commonly used in functionally guided radiotherapy [24]. In practice, a dose of 5 mCi of MAA is administered via injection of the vein (IV). These radiolabeled particles are able to lodge within small pulmonary arterioles and capillaries, reflecting the pulmonary perfusion level. Independent investigations also demonstrated a linear relation between perfusion SPECT-weighted radiation dose to lung and the incidence of post-radiation therapy pulmonary toxicity [40].

Lung perfusion imaging measures the blood circulation within the lung and is commonly used in the clinic to check the presence of impaired lung perfusion, such as blood clots or abnormal blood flow. Impaired lung perfusion usually signals respiratory diseases, including pulmonary embolism, pulmonary hypertension, and chronic obstructive disease (COPD) [41]. For these patients, perfusion imaging is considered as the initial imaging test choice for diagnosis. Perfusion imaging can also help to assess the impact of lung volume reduction therapy by looking into regional functional information [42]. For lung cancer patients, lung perfusion imaging can be utilized as a predictor of pulmonary function. Studies also showed perfusion images could guide FLART [35, 43].

Using SPECT V/Q, radiation treatment was shown to cause decreased perfusion of the irradiated lung tissue in a dose-dependent manner, with the most prominent effects occurring at radiation doses above 60 Gy. Decreased perfusion has been approved to be more sensitive in detecting lung injury after radiation therapy than decreased ventilation [44]. SPECT/CT perfusion/ventilation scanning is also used in radiation treatment planning and post-radiation follow-up for other non-pulmonary thoracic cancer [44]. In clinical practice, these scans are a helpful tool for follow-up and evaluation of lung injury after radiation [45-47]. However, there is a risk associated with a V/Q scan; even the amount of radiation exposure in this process is under the safe limitation. Problems can also arise due to radioactive substances and the insertion of intravenous therapy (IV). Patients may risk for excessive bleeding and infection at the IV site, allergic reaction to the radiation agents.
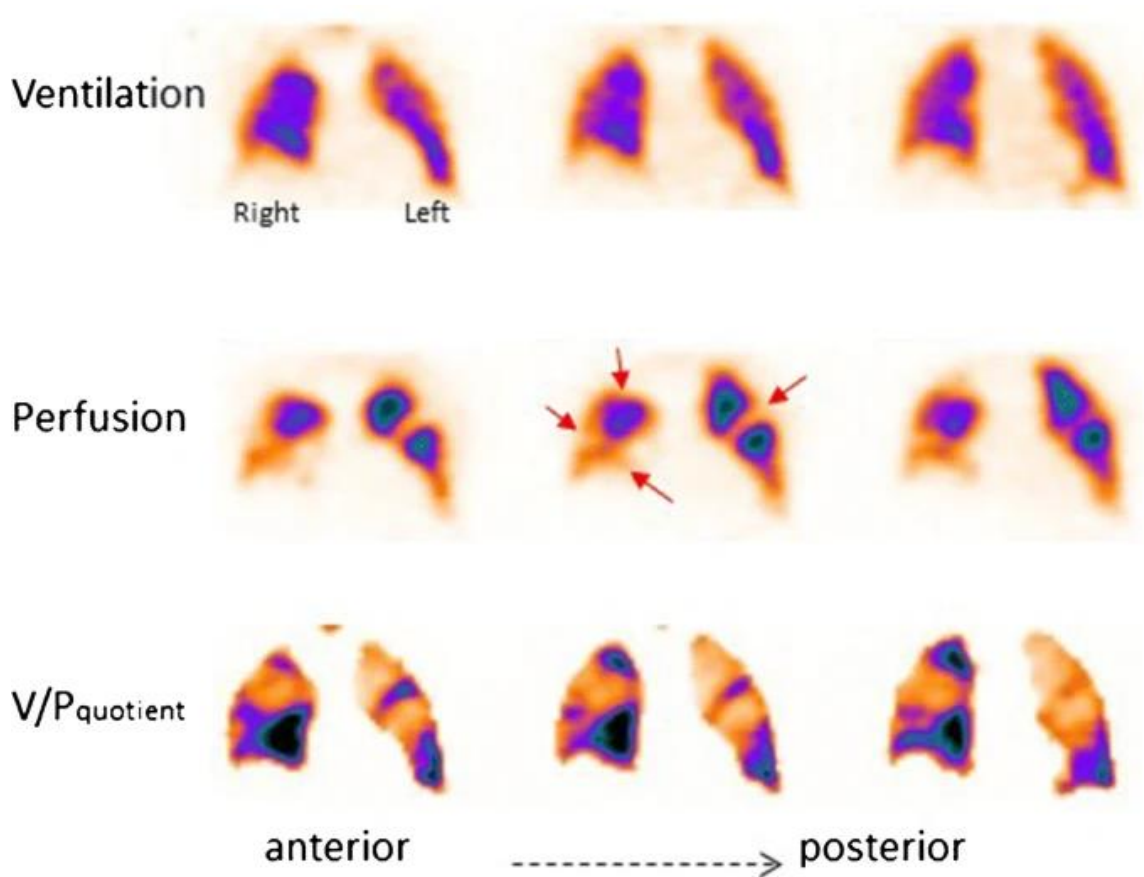
Figure 7 Illustration of ventilation and perfusion in coronal views. Red arrows indicate the multiple bilateral segmental perfusion defects [48].

1.2.3 Other Nuclear Pulmonary Function Imaging Methods

Over the past decades, lung function has traditionally been assessed using spirometry. By measuring the airflow volume and speed in inhaling and exhaling, the function of the lungs as a whole can be calculated [49]. However, the effects of disease on the lung are mostly restricted to regional lung volume, providing motivation to develop new methods that are capable of providing regional lung functional information. Besides SPECT, several other imaging modalities exist for pulmonary V/Q imaging [50, 51]. The representative methods are positron emission tomography (PET) with Ga-68 [52, 53] and hyperpolarized gas MRI [54-57] (Figure 8).

The use of MRI has been conventionally limited by the low signal in the lung since the low proton density in parenchyma tissue. To overcome these challenges, significant research has been dedicated to enhancing the contrast in the lung parenchyma. Oxygen-enhanced and fluorinated gas MRI were introduced in the early 1990s for lung ventilation weighted imaging [58, 59], but were not widely disseminated. Hyperpolarized (HP) gas MRI, particularly He-3 and more recently Xe-129 gases, have been demonstrated to provide a wider array of functional contrasts, including spin density, diffusion-weighted imaging, and gas exchange [60, 61]. Moreover, functional lung imaging can help us improve our understanding of a wide range of chronic lung diseases, including chronic obstructive pulmonary disease.

Compared with SPECT and MRI, PET can provide superior image quality and scan speed. In addition, PET is an inherently quantitative volumetric modality. Ga-68 is a positron-emitting radionuclide that has several convenient features for PET imaging: (i) half-life of 68 minutes; (ii) small on-site generator; (iii) chelation with a number of useful molecules [62]. Using the same carrier molecules of Technegas in SPECT V/Q

imaging, Ga-68 can be used to synthesize Ga-68 carbon nanoparticles, named "Galligas", for ventilation imaging. For perfusion imaging, Ga-68 can synthesize 68Ga-macro aggregated albumin (Ga-68-MAA) [63]. Besides, most commercially available PET scanners are combined with CT as PET/CT, enabling detailed anatomic correlation. Ga-68 PET offers a unique opportunity to further improve the accuracy and utility of V/Q imaging.

However, the clinical practice of nuclear medicine ventilation imaging is limited for a variety of drawbacks. SPECT ventilation imaging only offers a limited spatial resolution. PET needs a long imaging time. Ga-68 has ionizing radiation. MRI ventilation imaging needs hyperpolarized gas. These limitations make them difficult to implement. In general, these three modalities are low accessibility for the radiation oncology department and invasive techniques for the patients [53, 57].
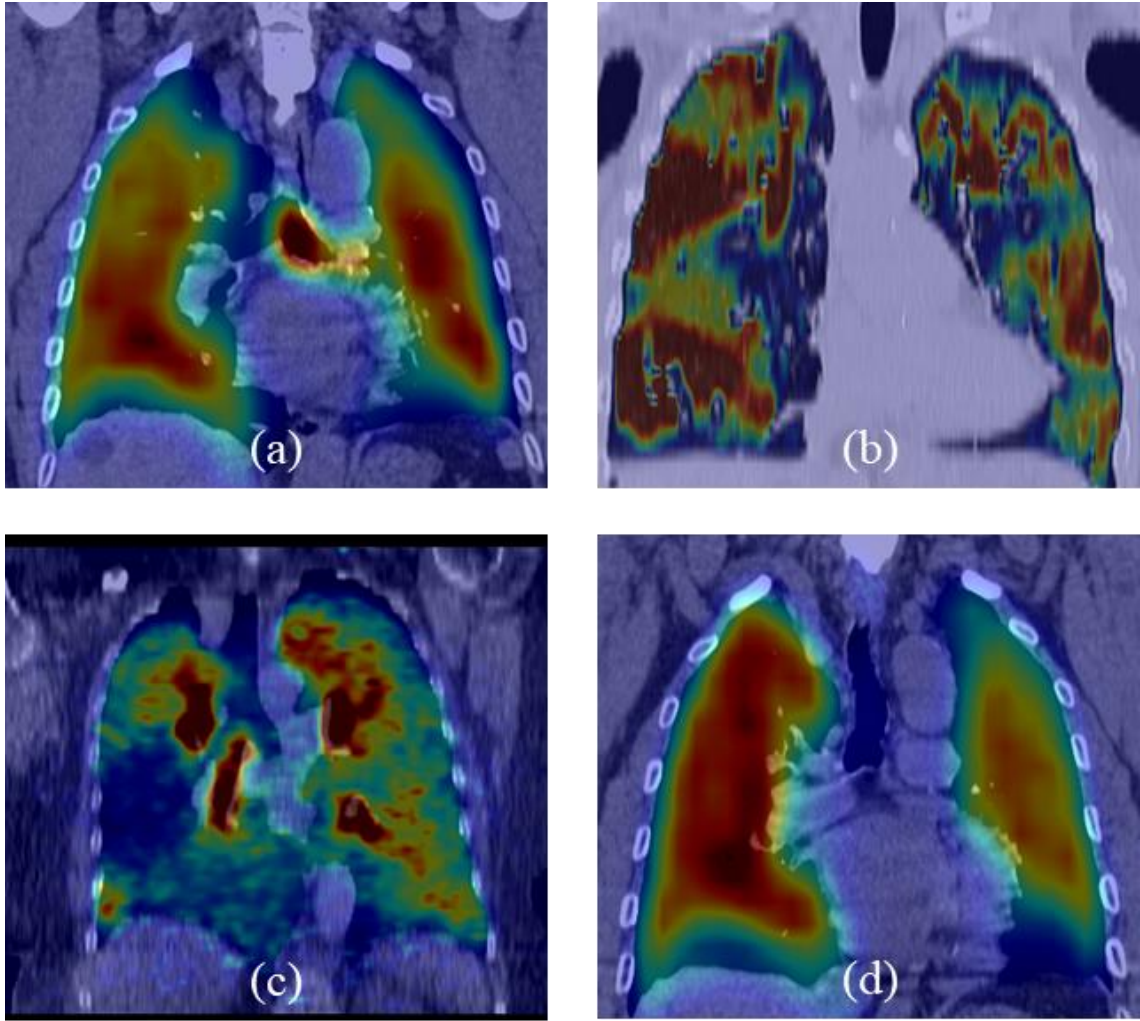
Figure 8 Examples of nuclear based ventilation/perfusion imaging. (a) DTPA SPECT perfusion; (b) Hyperpolarized gas MRI ventilation; (c) Galligas PET ventilation; (d) MAA SPECT perfusion.

1.2.4 CT Based Ventilation/Perfusion Imaging

Nuclear V/Q imaging is the current clinical standard of functional lung imaging. But the application has been limited for several reasons. To address this processing clinical need, several CT based imaging techniques have been under development. 4D-CT, which is used for CTVI calculation, is a standard of care for lung cancer radiotherapy, and deriving a functional map from 4D-CT involves image processing/analysis. Obtaining the lung functional images using the CT images were proposed as an alternative method for nuclear medicine imaging [64]. As a result, CT ventilation imaging (CTVI) was proposed. Recent efforts have focused on CT-derived ventilation imaging (CTVI) [65]

The typical CTVI uses deformable image registration (DIR) algorithm to detect the voxel movement in the lung region, which reflecting the lung density or volume changes between the exhale and inhale CT images [53]. Most CTVI approaches depend on two fundamental steps (Figure 9): 1. 4D image acquisition, 2. DIR between the inhale and exhale CT images [57]. Then, different ventilation metrics were used to process the DIR outcome and compare it with the radiation medicine images. For example, Reinhardt et al. used the Jacobian determinant motion field of DIR to estimate volume changes in the lung region [66]. Kipritidis et al. estimated regional lung volume density changes between the deformable vector filed between the exhale and inhale CT images [52].

CTVI has advantages over other modalities (SPECT, PET, and MRI) in the following two aspects: 1. CTVI has higher resolution, lower cost, shorter scan time, and greater availability; 2. CTVI does not bring in extra cost because CT is a standard of care for lung cancer radiotherapy. In general, CTVI presents exciting opportunities in

23

radiotherapy to account for functional lung status in creating individualized plans for patients [53, 57].

However, because CTVI is computed/synthesized from CT images, the accuracy of CTVI depends on the synthesizing techniques and the quality of input images. It is critically important to validate CTVI against more established modalities of ventilation (SPECT, PET, and hyperpolarized gas MRI). However, much research has proved that it is difficult to establish robust voxel-level correlations between CTVI and nuclear medicine functional map. The results of a recent international multi-institutional study on CTVI validation, the VAMPIRE (Ventilation And Medical Pulmonary Image Registration Evaluation) Challenge [57], showed that the correlation between CTVI and the reference modalities vary with the CTVI method, the reference modality, the imaging subject, and the evaluation metric. This concerning variability of CTVI highlights the importance of developing new, reliable, and accessible ventilation imaging technique for functionally guided radiotherapy applications.

Since CTVI is synthesized images from acquired CT images, the accuracy of CTVI ventilation images depends on the synthesizing techniques and the quality of input images. It is, therefore important to validate CTVI against more established nuclear medicine ventilation Imaging (SPECT, PET, and hyperpolarized gas MRI). However, DIR setting is experience-based because this algorithm has many parameters that need to set, and the DIR outcome is also sensitive to the CT artifacts. For these reasons, the DIR-based CTVI has large variance and limited accuracy depending on the tuning of the DIR algorithm. The DIR algorithm based CTVI only achieved a median of 0.49 correlation coefficient in the range of 0.27-0.73 [57].
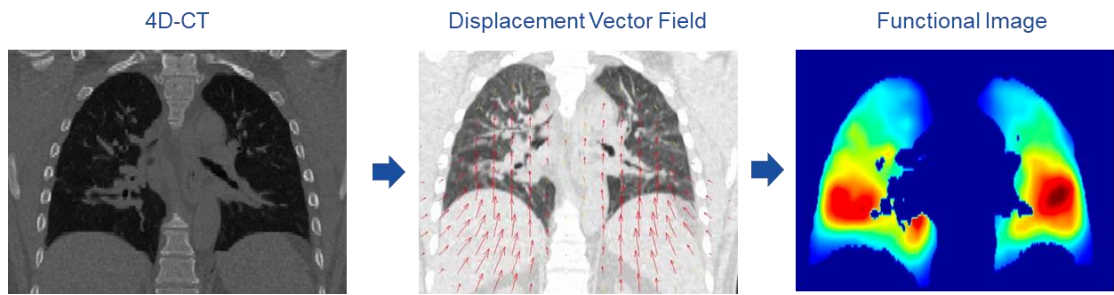
Figure 9 Illustration of the workflow for creating 4D-CT ventilation imaging methods.

1.3   *Deep Convolutional Neural Network*

1.3.1   Development of Artificial Intelligence Techniques

We have witnessed the overwhelming development of Artificial Intelligence (AI) techniques in the past decades. AI is a broad concept with varying definitions. In the early review, Simmons and Chappell define intelligence as an ability to resolve definite problems and an ability to study solutions for new problems [67]. In the latest review, Fazal further states AI to the ability to perceive its environment, search and perform pattern recognition, plan and execute an appropriate course of action and perform inductive reasoning to derive general principles [68]. As shown in Figure 10, the term AI can be dated back to the 1960s. After the development of the computer in science and industry, researchers tried to develop a system to automatically find diagnosing conditions from radiological images. Although interesting results were achieved, the application was ultimately unsuccessful as the commuting power was insufficient. In the 1980s, a number of machine learning techniques were brought out, including rule-based and case-based reasoning, Bayesian networks, hypertext algorithms, et al. [69].
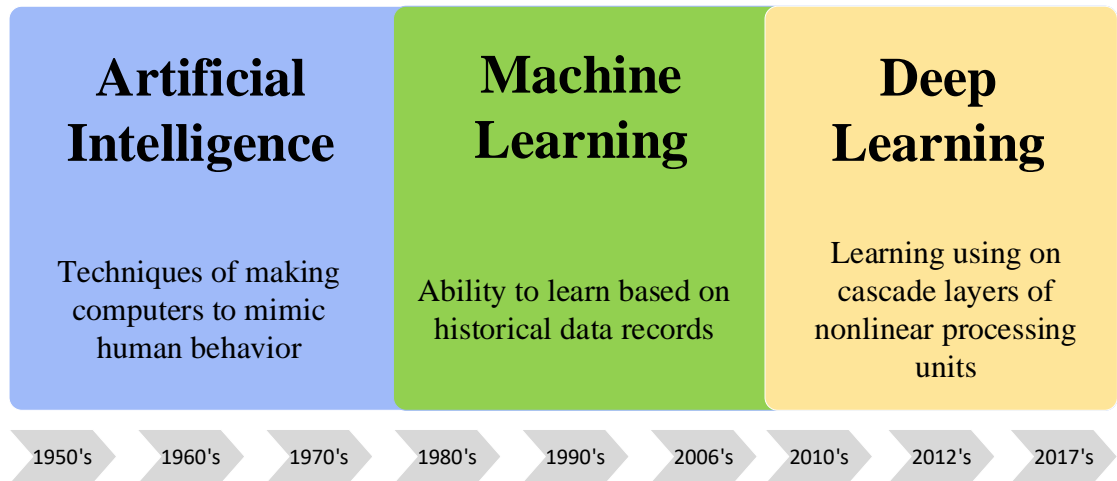
Figure 10 Development of deep learning techniques.

By far, the most successful techniques of AI system in medical imaging is the artificial neural network (ANN). ANN is featured by the function of simulating the way the human brain analysis and processes information [69]. An ANN is a collection of connected tunable units (also known as nodes, neurons, and artificial neurons) that can extract a hierarchy of features from input datasets and pass a signal from a unit to another.

In 1986, the term "deep learning" was first introduced to machine learning [70] and expected to overcome some limitations of classical machine learning-based methods. It is defined as a series of machine learning algorithms using several layers of nonlinear processing units for feature extraction and signal transformation. In 2000, Igor Aizenberg and colleagues used this concept for ANNs [71]. Although it is a broad concept, including stacked auto-encoders, deep Boltzmann machines, and ANN, it is usually referred to as ANN-based deep learning [72].

In deep learning, a deep convolutional neural network (CNN, or ConvNet) is the most commonly used architecture for image processing [73]. CNN is a class of ANNs characterized by one or more layers of convolution units. CNN is most commonly applied for image synthesis and gained great interest. CNN has consistently outperformed classical machine learning approaches [72] and achieved the impressive result on many medical segmentation challenges [74, 75] since 2012. These successful applications of the deep learning network make it a promising method. CNN is most commonly applied for image segmentation and gained great interest. after AlexNet won the ImageNet Large Scale Visual Recognition Competition [76]. Applied on daily-life object segmentation, CNN has nearly halved the error rates of the previously best segmentation approaches.

CNN can be distinguished from previously developed models in two aspects: (i) local connectivity. Each filter inside a convolutional layer is connected to a small region of the layer, named as receptive field. CNN can exploit spatial locality information by utilizing a local connectivity pattern between adjacent neurons. This structure ensures that the filters produce the strongest response to a region of spatially local input. Stacking many local connectivity layers leads increases the receptive regions so that the representations of small regions in input assembles representations of global areas. (ii) shared weights. Across the entire visual field of a CNN, each filter shares the same parameterization and forms a feature map. All the filters in each convolutional layer respond to the same feature in the response field. This method allows for the equivariant feature map under changing locations of input features in the entire visual field [73]. Local connectivity allows CNNs to achieve better generalization on image processing tasks. Weight sharing dramatically reduces the number of learned parameters, thus lowering the memory requirements for running the network and allowing the training of deep networks.
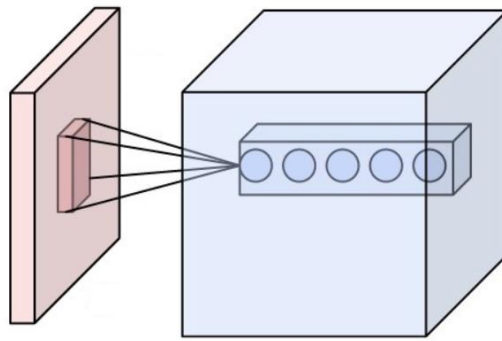
Figure 11 Illustrations of the neurons in a convolutional layer (blue), connected to

their receptive field (red) [73].

1.3.2 Artificial Intelligence in Medical Image Synthesis

Image synthesis, also known as image translation, refers to an image-to-image transformation in the context of radiation oncology. The pair of images for synthetization can be of different image modalities or acquired under different conditions. Owing to recent advances in deep learning, AI has achieved remarkable progress in its applications of image synthesis [77]. Below, we summarize the status of image synthesis in terms of applications and technology.

(1) *Inter-modality synthesis*. The supplementary information provided by synthetic images is beneficial to many radiotherapy applications, including positioning verification, dose calculation, and adaptive radiotherapy. Magnetic resonance (MR) images – known for their rich anatomical and functional information of tissues and organs – have been used to synthesize computed tomography (CT) images, namely MRI-to-CT conversion, for MRI-only radiation therapy treatment planning [78]. The reverse procedure, CT-to-MRI conversion, was also proposed to address the problem of misalignment between the two image modalities and improve soft-tissue visibility in CT images [79, 80]. For example, the synthetic MRI has been used for prostate segmentation in CT images [81]. In addition to volumetric translation, MR projection to CT projection translation was implemented to achieve higher soft-tissue contrast and faster imaging speed during interventional therapy [82]. Furthermore, lung functional imaging (perfusion, ventilation) synthetization based on CT images has recently been proposed for functional avoidance radiotherapy [83, 84].

(2) *Intra-modality synthesis*. The accessibility of high-performance scanners is usually limited due to various factors such as the patient's health condition, time consumption, and monetary cost. Intra-modality synthesis aims to improve the quality

of images acquired from low-performance scanners by enhancing contrast and reducing artefacts. For example, 3T-to-7T MR image translation has been proposed to facilitate diagnosis and prognosis through the mining of detailed soft-tissue information in 7T MR images [85]. Another example is cone-beam CT (CBCT)-to-CT translation, wherein image artefacts and scattering noise can be largely reduced, improving patient positioning accuracy in radiotherapy [86]. Generally speaking, AI-assisted image synthesis can provide better images for diagnosis and therapy while reducing cost and workload.

The most common deep learning model in AI-assisted image synthesis is the deep convolutional neural network (DCNN), in which an encoding process extracts features and another decoding process utilizes those features for synthesis [78, 83, 84]. Recently, generative adversarial network (GAN)-based models attracted great attention for their unique adversarial loss and better performance [79, 86, 87]. Then, GAN derivatives, including Cycle GAN and MR-GAN, have been used to provide unsupervised learning where exact matching anatomy is not required [79, 80].

1.3.3 Classic Convolutional Neural Network Models

CNN models have achieved overwhelming advantages over conventional machine learning algorithms for image processing tasks. In this section, several classic CNN models were reviewed.

Fully convolutional network (FCN): FCN was brought out in 2014 for the task of semantic segmentation [88]. FCN is composed of convolutional layers without any fully connected layers throughout the network (Figure 12). Compared with previous AI models, the main difference is that the FCN is learning filters everywhere. A conventional CNN with fully connected layers is just as end-to-end learnable as a fully convolutional one. For FCN, even the end decision layers are learnable filters. An FCN tries to learn representative features and make a prediction based on local input information. In the end, the fully connected layer learns the global feature information where the local arrangement of the input is ignored. It can efficiently learn to make a dense prediction for per-pixel tasks. Unlike the previously introduced CNNs, an FCN transforms the output back to the original size of the input image through a transposed convolutional layer. Besides, FCN has the following features: (i) upsampling via transpose convolution. Convolution is a mathematical operation to make the size of the output feature map smaller. Thus transpose convolution is the revision process to get the size of the output feature map larger. It should note that transpose convolution, also named up convolution, deconvolution, and fractional stride convolution, when stride convolution is used. (ii) Fusing the output. Deep features can be obtained when the network structure goes deeper. However, spatial location information is also lost in deeper layers. FCN combines the output of shallow and deep layers to enhance the location information and deep features.

33
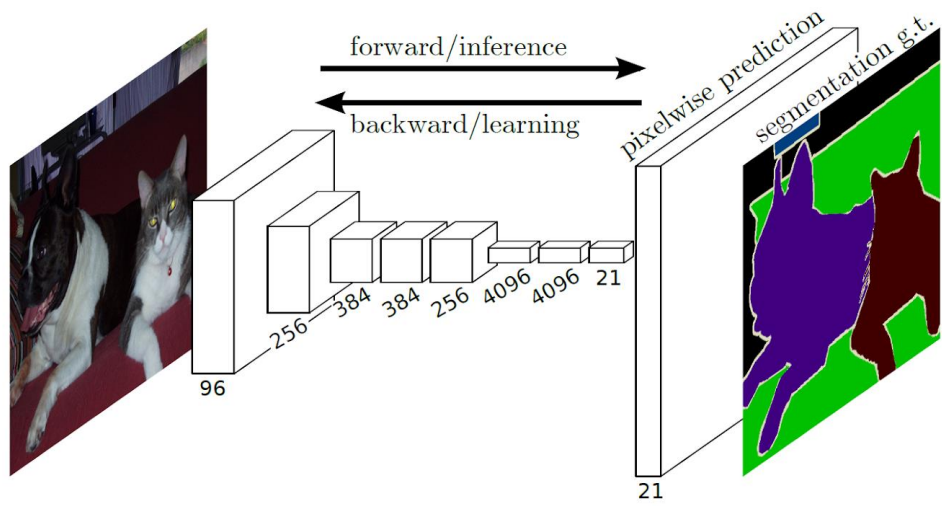
Figure 12 Architecture of the FCN

U-Net: U-Net was firstly developed in 2014 for the segmentation of neuronal structures in electron microscopic stacks and yield more precise segmentation than previous models [89]. The U-Net architecture is grounded on the FCN architecture [88]. U-Net got its name from the u-shaped architecture (Figure 13), which consists of a contracting path and an expansive path. The contracting path is a typical convolutional network that consists of repeated layers of convolutions. Each convolution is followed by activation function of rectified linear unit (ReLU) function and a max-pooling operation. During the contraction, the spatial information is reduced while feature information is increased. The expansive pathway combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the contracting path. Long skip connections were formed in a symmetrical manner to recover the fine-grained details in the prediction.

There are many applications of U-Net in the field of medical image processing, such as organ segmentation and inter-modality image synthesis. Different variations of the U-Net have also been widely applied for different medical image tasks. Here are some variants and applications of U-Net as follows: (i) 3D U-Net. This model is an updated U-Net from 2D to 3D for dense volumetric segmentation from Sparse Annotation [90]. The entire 3D images can be used as the input to the model. (ii) V-Net. This model has several improvements from U-Net. It introduces residual blocks and uses element-wise for horizontal residual connection. Besides, the convolutional layers replaces the up-sampling and down-sampling layers. [75]. (iii) Attention U-Net. This architecture learns where to look for features with the help of attention modules for the pancreas segmentation [91].
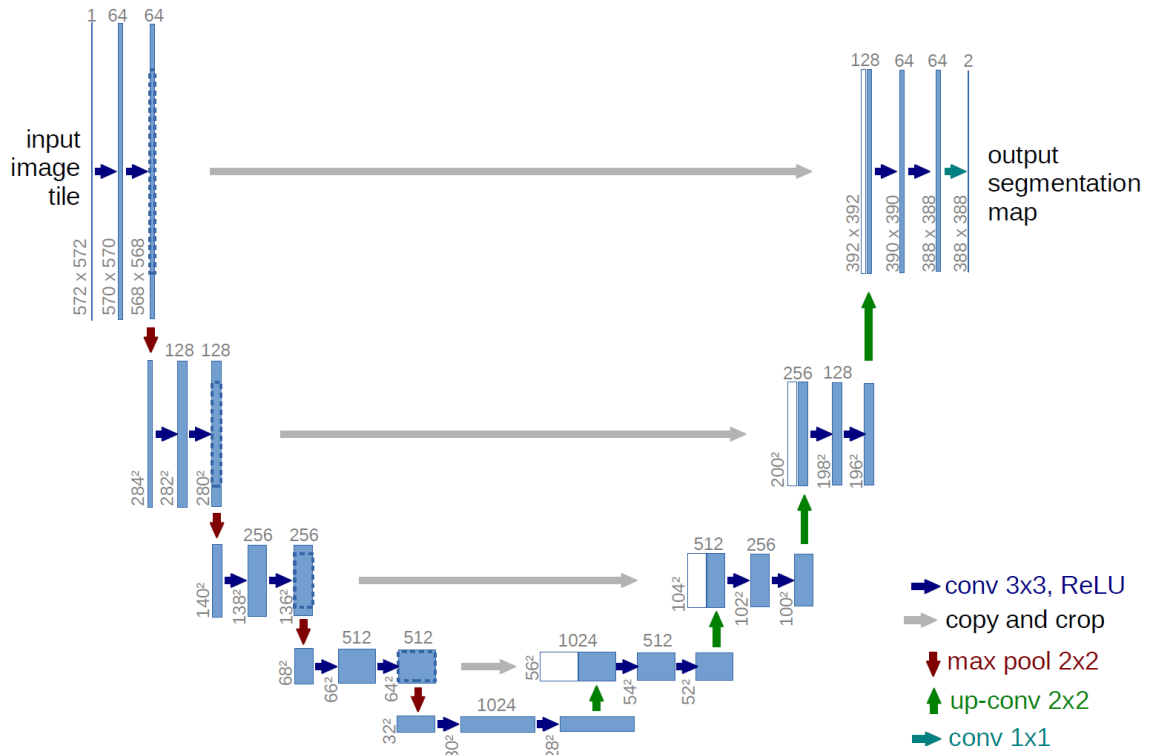
Figure 13 Architecture of the U-Net [89].

Generative adversarial networks (GANs): This category of networks is a framework for teaching a deep learning model to capture the features of the training data via an adversarial process so new data can be synthesized using the same features. GANs was firstly described in 2014 [92]. Each GAN is composed of two distinct models: a generator that is trained to generate fake images; a discriminator that tries to classify an image as real or fake (Figure 14). During training, the generator is constantly trying to fool the discriminator by generating better fakes, while the discriminator is working to classify the real and fake images with higher accuracy. The equilibrium of this game is when the generator is generating perfect fake images that look as if they are picked out from the training set, and the discriminator is left to always guess at 50% confidence that the generator output is real or fake.

To fully explore the potential application in image processing, many variants of GANs have been developed. Deep convolutional GAN (DCGAN) is a direct extension of the GAN described above for unsupervised learning, except that it explicitly uses convolutional and convolutional-transpose layers in the discriminator and generator, respectively [93]. Another important category is conditional GAN (cGAN), which learns a mapping from some conditional settings to images. The representative model is Pix2Pix GAN. The generator of a Pix2Pix GAN (Figure 14) is a convolutional network with U-Net architecture. The input image will pass through a series of convolutional layers and produces an output image that is of the same size as the input. After the generator, the synthetic image is concatenated with the input image. This concatenated set is fed as the input of the discriminator network.
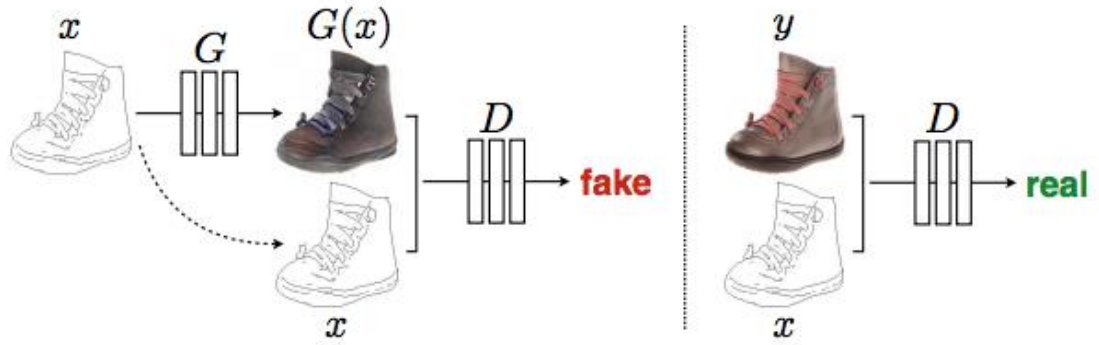
Figure 14 Architecture of the Pix2Pix GAN [93].

2.   Research Aims and Objectives

2.1   Research Gaps

Each of the developed pulmonary function imaging methods has its merits and drawbacks. SPECT ventilation/perfusion is the most widely used method but involves ionizing radiation and suffers from relatively low spatial and temporal resolutions. PET ventilation provides more quantitative information but requires extra ionizing radiation and a long imaging time. Hyperpolarized gas MRI ventilation produces exquisite images of pulmonary ventilation with superior spatial resolution and is free of ionizing radiation. However, the limited availability of the Helium-3 gas and the need for specialized equipment impedes its clinical implementations. Contrast-enhanced MRI/CT perfusion requires the injection of imaging contrast agents, and the data processing procedures can be complicated. 4D-CT based ventilation is error-prone due to un-robust DIR algorithms. In general, all these methods require additional scan(s) that are resource-demanding, inconvenient, and technically challenging, thus reducing the availability of personalized lung functional imaging for diagnosis and treatment.

To overcome these drawbacks of the current pulmonary function imaging techniques, we proposed to synthesize pulmonary functional images based on the textural information from CT images. In particular, we aim to develop a deep learning-based CT function mapping (DL-CTFM) method for the pulmonary CT-to-perfusion translation; and explore the feasibility of CT-to-ventilation translation using transfer learning. In clinical practice, lung perfusion imaging is a commonly used clinical pulmonary function test that measures the blood circulation within the lungs. It is commonly utilized to detect abnormal blood flow to assist in the diagnosis of respiratory diseases characterized by impaired lung perfusion function [43]. Deep

learning-based convolutional neural networks (DL-CNNs), also known as deep neural networks, are a class of machine learning algorithms characterized by multi-layers of convolution units. This technique has gained increasing popularity in medical imaging research and demonstrated satisfactory performances in many applications of medical image analysis, such as low-dose CT image correction [94, 95], MR-to-CT image synthesis [79], and image segmentation [96, 97]. These applications raise the expectations that a deep neural network can extract textural information from CT images and generate the corresponding pulmonary functional images.

2.2 Research Aims

This research aims to overcome the aforementioned deficiencies of current established pulmonary function imaging techniques by developing an artificial intelligence-based functional lung imaging methodology. We achieved this goal by developing, optimizing, and logically synergizing a number of novel image processing and synthetic techniques. Our overall rationale is that with the significantly improved function imaging consistency and imaging quality, the functional images will enable us to more accurately protect normal lung tissue and design better functional lung avoidance treatment plans for lung cancer patients. Once fully developed, this end-to-end functional imaging framework can be applied to any type of CT technique to develop functional images regardless of the image acquisition and reconstruction methods.

2.3 Research Objectives

1. Develop a deep learning-based CT function mapping (CTFM) framework for pulmonary CT-to-perfusion translation in a retrospective study

2. Validate the potential application of the developed model for functionally guided lung cancer radiotherapy

3. Explore the feasibility of CT-to-ventilation translation using a transfer learning strategy

Compared to current methods of pulmonary function imaging, our method is easier, faster, less costly, and more accessible for clinical implementation of functionally guided radiotherapy. Owning to these unique advantages, our method holds great promises to expedite and expand the clinical adaption of functionally guided radiotherapy. Ultimately, it may improve the treatment outcome of lung cancer radiotherapy by reducing radiation-induced lung injury and benefit lung cancer patients worldwide.

3.   Materials and Methods

3.1   Image Acquisition and Patient Characteristics

3.1.1   QMH Perfusion Dataset

SPECT/CT perfusion images of 75 patients who underwent lung MAA SPECT/CT scan for suspected lung disease and treatment evaluation in 2019 at Hong Kong Queen Mary Hospital were retrospectively collected. Two patients were excluded due to incomplete lung images. Among the 73 patients, 12 patients were diagnosed with lung cancer, and four patients were diagnosed with no disease. The remaining patients had different types of lung diseases such as pulmonary hypertension, chronic pulmonary embolism, and pulmonary arteriovenous malformation. The characteristics and diseases of patients are listed in Table 1. Patients were injected with 3 mCi (111 MBq) radiopharmaceutical

MAA before imaging and immobilized in the supine position with normal resting breathing during image acquisition. The 3D SPECT/CT scans were acquired in 360 degrees to cover the whole lung volume by the GE Discovery 670 SPECT/CT scanner with a frame rate of 30 s/f and a total frame number of 60. Each CT image was reconstructed into a 512×512 matrix with 0.977×0.977 mm^2 pixel spacing and 1.25 mm slice spacing, and each SPECT image was reconstructed into a 128×128×128 matrix with 4.42×4.42×4.42 mm^3 voxel size.

Table 1 Patient characteristics of QMH perfusion dataset (n=73)

		Number	**Percent**
Sex	Male	24	32.9%
	Female	49	67.1%
Age	Mean ± SD	67 ± 14	
Disease	Pulmonary hypertension	25	34.2%
	Pulmonary embolism	12	16.4%
	Lung cancer	12	16.4%
	Chest pain	2	2.7%
	Chronic obstructive airway disease	2	2.7%
	Systemic lupus erythematosus	2	2.7%
	Others*	14	19.2%
Non-Disease	Healthy lung	4	5.5%

*Others: Pulmonary fibrosis, hypoxia, congestive heart failure, aortic dissection, pulmonary arteriovenous malformation, chronic renal failure, hepatopulmonary syndrome, fever, post COVID-19, shortness of breath, hypoplastic right ling, tuberculosis, pneumonia, deep venous thrombosis.

3.1.2 QMH Ventilation Dataset

SPECT/CT ventilation images of 13 patients who underwent lung Technegas SPECT/CT scan for suspected lung disease and treatment evaluation in 2019 at Hong Kong Queen Mary Hospital were retrospectively collected. Techegas was administered prior to acquiring the ventilation SPECT. Images were acquired with patients in slow and deep breaths to maximize dispersal of aerosol in the pulmonary parenchyma. Patient characteristics are listed in Table 2. The 3D SPECT/CT scans were acquired in 360 degrees to cover the whole lung volume by the GE Discovery 670 SPECT/CT scanner with a frame rate of 30 s/f and a total frame number of 60. Each CT image was reconstructed into a 512×512 matrix with 0.977×0.977 mm^2 pixel spacing and 1.25 mm slice spacing, and each SPECT image was reconstructed into a $128 \times 128 \times 128$ matrix with $4.42 \times 4.42 \times 4.42$ mm^3 voxel size.

Table 2 Patient characteristics of QMH ventilation dataset (n=13)

		Number	Percent
Sex	Male	1	7.7%
	Female	12	92.3%
Age	Mean ± SD	66 ± 13	
Diagnosis	Pulmonary hypertension	7	53.8%
	Pulmonary embolism	3	23.1%
	Lung carcinoma	1	7.7%
	Tuberculosis	1	7.7%
	Acute on chronic renal failure	1	7.7%

The Institutional Review Boards (IRB) of the University of Hong Kong/Hospital Authority Hong Kong West Cluster.

3.2 *Deep Learning-based CT Function Mapping (DL-CTFM) Framework*

We developed a deep learning-based CT function mapping (DL-CTFM) framework for pulmonary perfusion mapping. This framework is able to synthesize pulmonary perfusion images from CT images. It is composed of three parts: This framework includes three parts: image preparation, image processing, and our proposed convolutional neural network (CNN). Image preparation consists of a series of morphological operations to decrease computational consumption and remove noise. Image processing aims to enhance the robustness of CT image features and standardize SPECT perfusion images to be suitable labels for CNN mapping. It consists of SPECT normalization, SPECT discretization, CT contrast enhancement, and CT defect enhancement. The proposed CNN was constructed to extract features from the processed CT image and synthesize the perfusion image.

3.2.1 Image Preparation

The image preprocessing procedures are illustrated in Figure 15 Each SPECT image was registered to the CT geometry using rigid transformation in 3D Slicer (10.4.2) [98]. A parenchyma mask was automatically generated in CT images using the Chest Imaging Platform (2019), which is an open-source library and workstation for quantitative chest imaging [99]. The trachea and tumor regions were excluded from the mask. This mask was subsequently applied to the SPECT and CT images to segment the parenchyma volume. In order to increase the size of the dataset and reduce the computational cost, the left and right lungs were separated. To further speed up the training process and reduce computational memory cost, the separated parenchyma

volumes were cropped to include only the lung and down-sampled to 128×64×64-sized matrixes.

3.2.2 Image Processing Pipeline

Image processing aims to enhance the robustness of CT extracted features and standardize SPECT perfusion images to be suitable labels for CNN mapping. The down-sampled CT and SPECT images were standardized using CT enhancement, SPECT normalization/discretization, respectively, as described below.

CT enhancement: CT enhancement is achieved through histogram equalization with 2 filters applied. This histogram equalization has been demonstrated to be effective in displaying different regions of the CT image in greyscale [100]. To enhance the contrast in the parenchyma region, the histogram of the down-sampled CT images was equalized within -1000 to -300 Hounsfield Unit (HU) values. The application of filters has been proven to improve the correlation between CT and lung functional images [101]. A median filter with a size of 10 voxels was subsequently applied to enhance the signal in low functional regions and to reduce the noise fluctuation. This was followed by a uniform filter with a kernel size of 5 pixels for further noise reduction. This CT enhancement procedure, including histogram equalization and filtering, was summarized by Equation 1:

$$CT(x,y,z)' = filtering(\frac{chf[CT(x,y,z)] - chf_{-1000}}{chf_{-300} - chf_{-1000}}), \qquad 1$$

where $CT(x, y, z)$ is the original HU value at voxel (x, y, z) and $CT(x, y, z)'$ indicates the enhanced CT image. Cumulative histogram function (chf) gives the cumulative count for a single HU bin. chf_{-1000} is the chf for the voxel intensity of -1000 and chf_{-300} for -300. Outlier pixel values were replaced with the threshold values. $filtering$ indicates the deployed filters.

SPECT Normalization and Discretization: The pulmonary SPECT value can be influenced by many factors depending on patients' conditions, such as respiratory capacity, breathing frequency, and lung diseases. SPECT normalization was commonly used to reduce the impact of individual patient conditions during inter-patient lung functions assessment [101-103]. In this study, a SPECT image was first normalized by the perfusion value of the high functional region. The normalized image was used as the ground truth for model evaluation. It was subsequently discretized into 11 uniformly spaced values between 0 and 1 (0, 0.1, 0.2, … 1) for modeling. Image discretization decreases the noise-induced variance of the extracted features and speeds up the modeling convergence. It has been previously adopted in PET ventilation images to mitigate the impact from small noise fluctuations [53]. The SPECT standardization procedure, including normalization and discretization, was formulated by Equation 2:

$$SPECT(x, y, z)' = discretize \left(\frac{SPECT(x, y, z)}{HIR} \right),$$
$$HIR = P_{75}(SPECT) \qquad\qquad 2$$

where $SPECT(x, y, z)$ is the original SPECT value at voxel (x, y, z) and $SPECT(x, y, z)'$ indicates the standardized SPECT image. HIR is the perfusion value of the high function region. In this study, it was set as the 75th percentile value for each SPECT $(P_{75}(SPECT))$ as it is close to the perfusion value of the normal tissue [104]. $discretize$ indicates the deployed discretization.
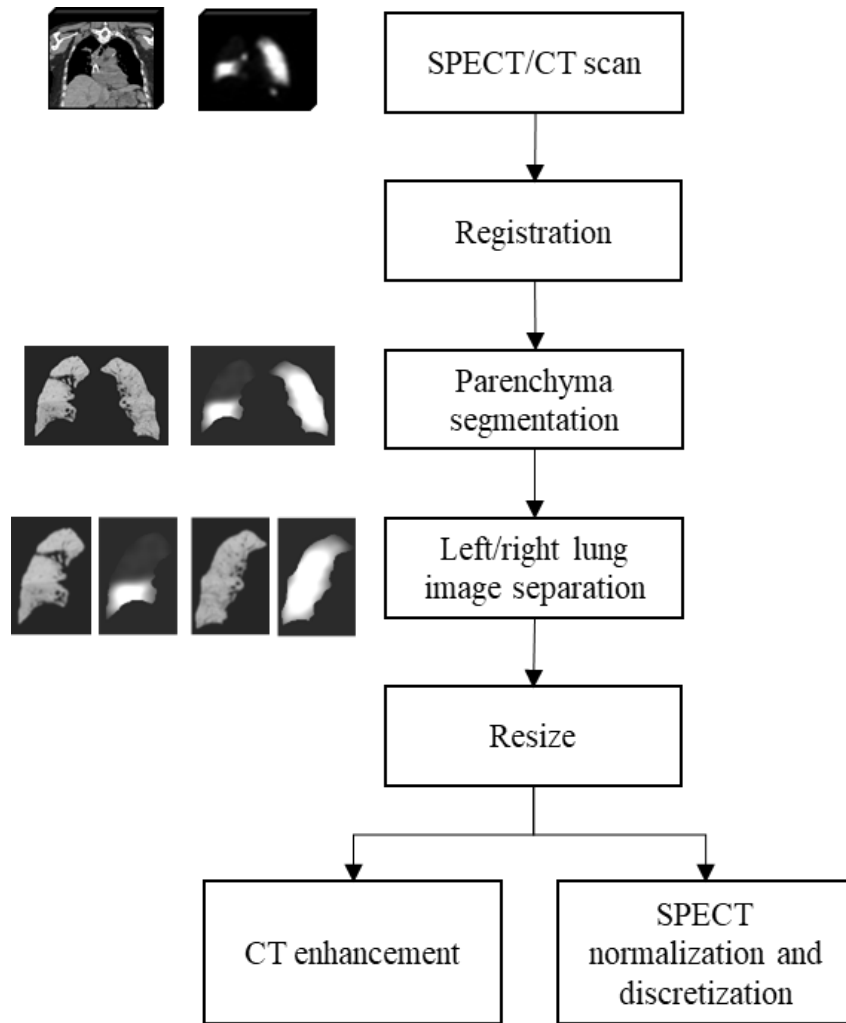
Figure 15 Flowchart (right) and images (left) of the image preparation and processing

pipeline.

3.2.3 Attention Residual Neural Network (ARNN)

CT-to-perfusion translation requires a full exploration of high-level features on 3D images, which can only be achieved by a 3D CNN [105]. However, the training process of a conventional 3D CNN requires significant memory and computational power. As such, we proposed an attention residual neural network (ARNN) with full utilization of a reduced number of parameters for the CT-to-perfusion translation.

This model used a 3D encoding-decoding structure to capture the hierarchical texture features of the input, as shown in Figure 16. The processed CT image is input at the beginning. Then the data is propagated through the network along all possible paths from left to right. In the end, the generated perfusion image comes out. The maximum number of channels was set to 128 to reduce the computational memory cost. Like other CNN models, ARNN includes a number of different modules. two skip attentions (Skip Att) (red blocks) similar to attention U-Net [91] were applied for target structure attention and original shape recovery [89]. Similar to ResNet [106], residual block [106] (ResConv) was adopted to speed up the convergence; strided convolution (StriConv) (purple bars) and transposed convolution (ConT) (light green bars) were used to half the feature map size (S1 to S4) and recover the feature map size, respectively. In order to increase the receptive field of the 3D architecture, we adopted three $2\times2\times2$ strided convolution modules and utilized other filters with $5\times5\times5$ filter size. Compared with the commonly used $3\times3\times3$ convolution filters, the $5\times5\times5$-sized filter can increase the volume of the receptive field by a factor of approximately $5^3/3^3$ = 4.6.
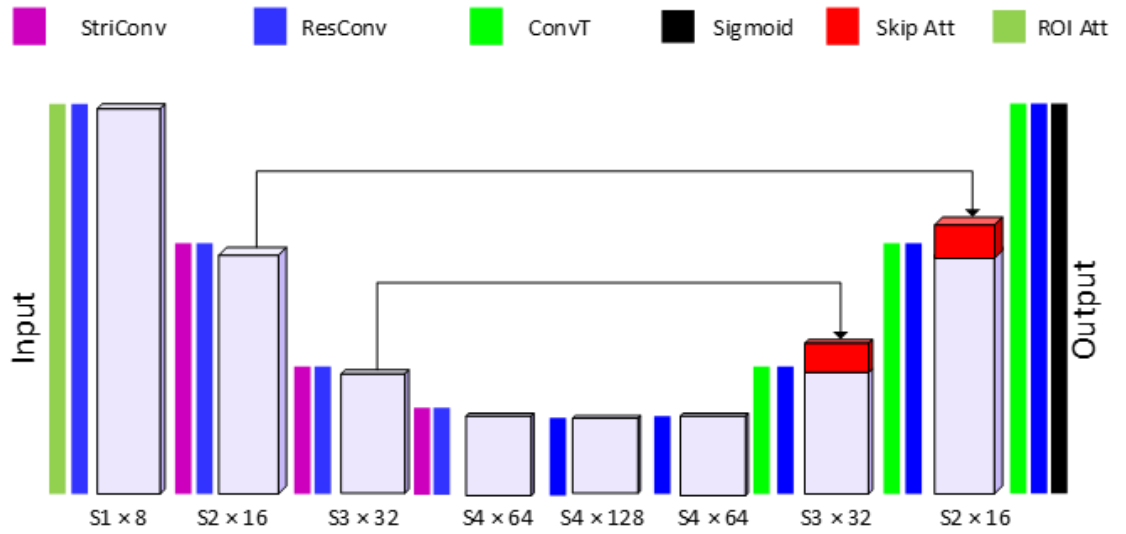
Figure 16 The architecture of the proposed CNN model. The 3D blocks indicate the feature map. StrideConv was short for stride convolution. ResConv is short for residual convolution. ConvT is short for convolution transpose. Sigmoid is short for the sigmoid layer. Skip Att is short for the skip attention module. ROI Att is short for ROI attention.

The region of interest (ROI) attention module (ROI Att) was also designed for foreground region attention (Figure 17-c). To obtain the ROI attention feature map, the input feature map was convoluted to increase the feature channel to 4, followed by thresholding to by setting each pixel to 1 if it is larger than drop threshold 0. After a followed sigmoid layer, this binary mask is applied to the convoluted feature map by spatial wise multiplication to obtain the attention feature map. In the last layer, a Sigmoid function (Sigmoid) sums up the results of the previous layers and scales the predicted values to a range of [0,1]. Each convolution is followed by batch normalization, a parametric rectified linear unit (PReLU) with rectification parameter of 0.2, and a dropout layer with a dropout rate of 0.1.
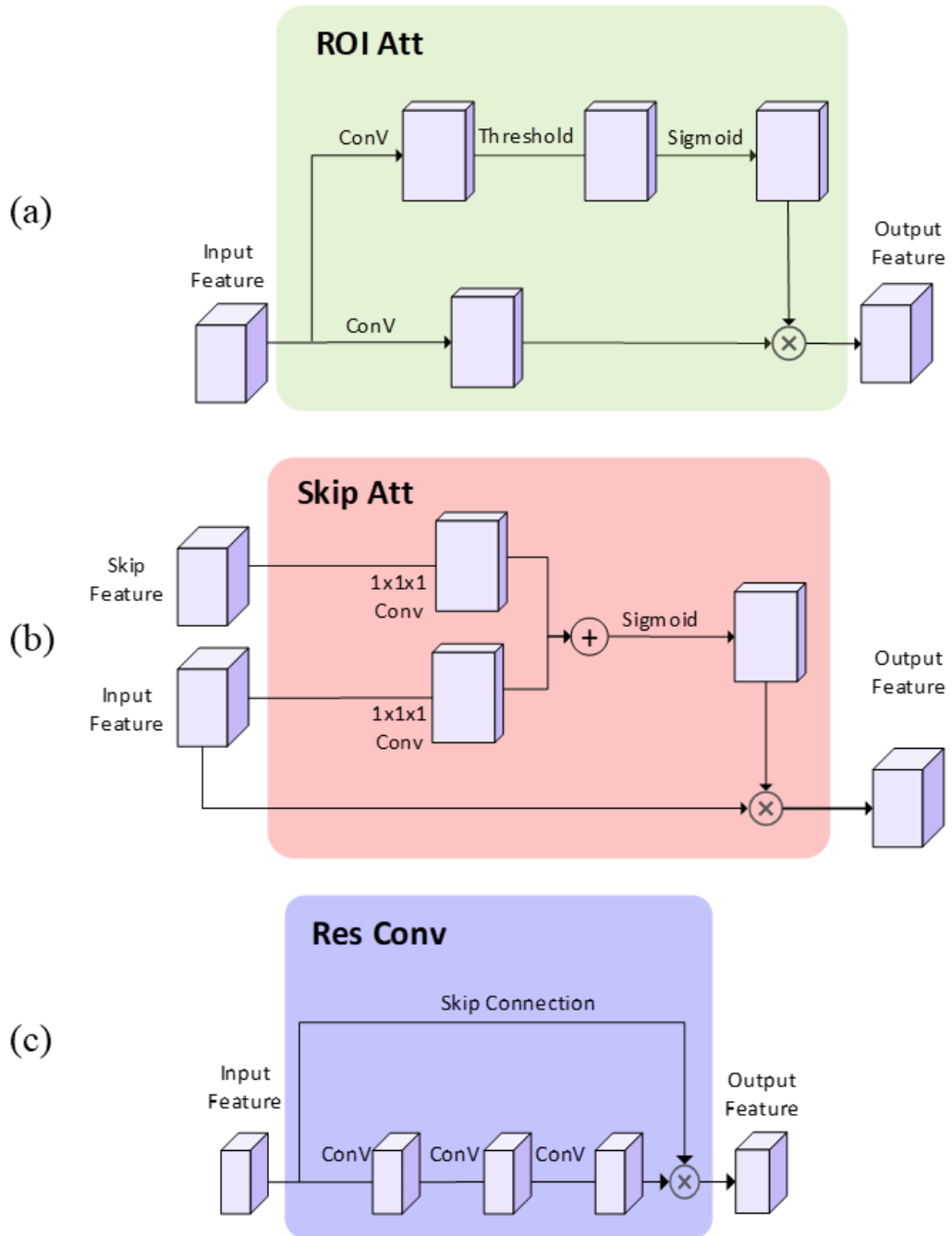
Figure 17 StrideConv was short for stride convolution. ResConv is short for residual convolution. ConvT is short for convolution transpose. Sigmoid is short for the sigmoid layer. Skip Att is short for skip attention. ROI Att is short for ROI attention.

3.3   *Investigation of the DL-CTFM Framework*

The impacts of three components (residual module, ROI attention, and skip attention) were tested using ablation experiments. Residual modules have shown to be capable of speeding up the convergence [106]. In our model, eight residual modules were designed in the middle of the network. The ROI attention module was configured on the first layer to steer the data learning process focusing on the parenchyma volumes, avoiding degraded coverage of loss function owing to the meaningless learning on the image background (i.e., non-parenchyma region). The skip attention module has been found to benefit target structure attention[89]. Because low functional regions normally are smaller than functional regions, two skip attention connections were used to help focus on the low functional region.

3.3.1   Dataset and Study Design

In this study, 42 pulmonary perfusion MAA SPECT/CT scans were randomly selected from the QMH perfusion dataset. These patients were randomly split into a training group (31 patients) and a testing group (11 patients). The CT and SPECT images of each patient went through image preparation and image processing prior to model training and testing. In image preparation, the left and right lungs were separated to augment the dataset size to 84 (62 for training and 22 for testing). The effectiveness of different components and parameters of this framework were analyzed using a series of ablation experiments (Figure 18). The CTFM generated perfusion images were compared with the processed SPECT perfusion label images using the structural similarity (SSIM) and correlation coefficient (R). After deciding the final framework, the generated CTFM perfusion images were recovered to the original size with post-

processing and compared with the SPECT perfusion in original size using the Dice similarity coefficient (DSC) in clinical relevance.

3.3.2 CNN Configuration

Multiple elements were particularly designed to minimize the risk of model overfitting. Our proposed model uses dropout layers and early stop to alleviate the likelihood of overfitting. Each convolution and activation function are followed by a dropout layer with a dropout rate in the range of [0, 0.1, 0.2, 0.3, 0.4, 0.5].

Although a large filter size can facilitate global texture feature extraction in the receptive field, it simultaneously increases the number of parameters and decreases the learning ability of the model. In order to explore the effect of filter size on the learning ability, 5×5×5 convolution filters, and 3×3×3 convolution filters were tested. The number of CNN parameters was controlled by the CNN width. In this study, a series of CNN widths (3×32, 4×32, and 5×32) were also tested to optimize the fitting power of the CNN model.

ReLU activation functions have been proven to speed up the training process compared to classic sigmoid alternatives. In this study, a parametric ReLU (PReLU) [107] was also developed from ReLU, as formulated in Equation 3:

$$PReLU(x) = \begin{cases} x, & x > 0 \\ ax, & x \leq 0 \end{cases} \qquad 3$$

where a is a learnable parameter. This loss function does not zero out the negative input so that a small gradient is allowed when the unit is not active. For the encoding and decoding structure, the PReLU was used after each convolution layer. LeakyReLU

is PReLU with a fixed slop. The performances of PReLU, LeakyReLU (slope = 0.5), and ReLU were tested in the validation group.

3.3.3 Loss Function and Optimizer

The proposed method learns the optimal parameter values by minimizing the loss function between the label and the output of the network. During optimization, each layer was updated using error backpropagation with an adaptive moment estimation optimizer (ADAM) or SGD, AdaGrad. ADAM is a first-order gradient-based algorithm, designed for the optimization of stochastic objective functions with adaptive weight up with adaptive weight updates based on lower-order moments. We also investigated the impact of different loss functions. In this study, the performances using mean square error (MSE) loss, binary cross entropy (BCE) loss, and auxiliary loss of Spearman's correlation coefficient (R) and structural similarity ($SSIM$) as loss functions were envaulted.

We used binary cross-entropy loss as the loss function to increase the sensitivity to the small low-functional regions, as follows (Equation (3)):

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \cdot \log(p_i) + (1 - y_i) \cdot log(1 - p_i)], \qquad (3)$$

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - p_i)^2 \qquad\qquad 10$$

where p_i and y_i denote the predicted and ground-truth perfusion values at voxel i, respectively. N denotes the total number of non-zero voxels.

3.3.4 Evaluation Metrics

The framework performance of each experiment was evaluated using an analysis metric between the predicted perfusion and SPECT perfusion. The analysis metric was the sum-up of Spearman's correlation coefficient (R) to assess the intensity correlation (Equation (4)) and the structural similarity index measure (*SSIM*) to determine the perceptual similarity (Equation 5). The generated DL-CTFM images were up-sampled to the original SPECT resolution and compared with the normalized SPECT perfusion. R is defined by the following equation (Equation (4)):

$$R = \frac{\sum_{i=1}^{N}[(y_i - \bar{y}) \cdot (p_i - \bar{p})]}{\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{N}(p_i - \bar{p})^2}}, \tag{4}$$

where p_i and y_i denote the predicted and ground-truth perfusion values at voxel i, respectively. N denotes the total number of non-zero voxels, while $\bar{p}$ and $\bar{y}$ denote the average predicted and ground-truth perfusion values, respectively.

The *SSIM* contains three terms that compare three measurements of the generated perfusion and perfusion label: the luminance term ($L(y, p)$), the contrast term ($C(y, p)$), and the structural term ($S(y, p)$) [36]. The overall *SSIM* is a multiplicative combination of these terms, as follows (Equation (5)):

$$SSIM = L(y,p) \cdot C(y,p) \cdot S(y,p) = \frac{2\mu_y\mu_p + C_1}{\mu_y{}^2 + \mu_p{}^2 + C_1} \cdot \frac{2\sigma_y\sigma_p + C_2}{\sigma_y^2 + \sigma_p^2 + C_2} \cdot \frac{2\sigma_{yp} + C_2}{2\sigma_y\sigma_p + C_2}, (5)$$

where μ_y, μ_p, σ_y, σ_p, and σ_{yp} are the respective local means, standard deviations, and cross-variance for the ground-truth image y and predicted image p. $C_1 = (k_1L)^2$, $C_2 = (k_2L)^2$ are two variables that stabilize the divisions when the denominators are too small. L is the dynamic range of the voxel values. k_1 and k_2 are scalar constants, which are by default set to 0.01 and 0.03, respectively.
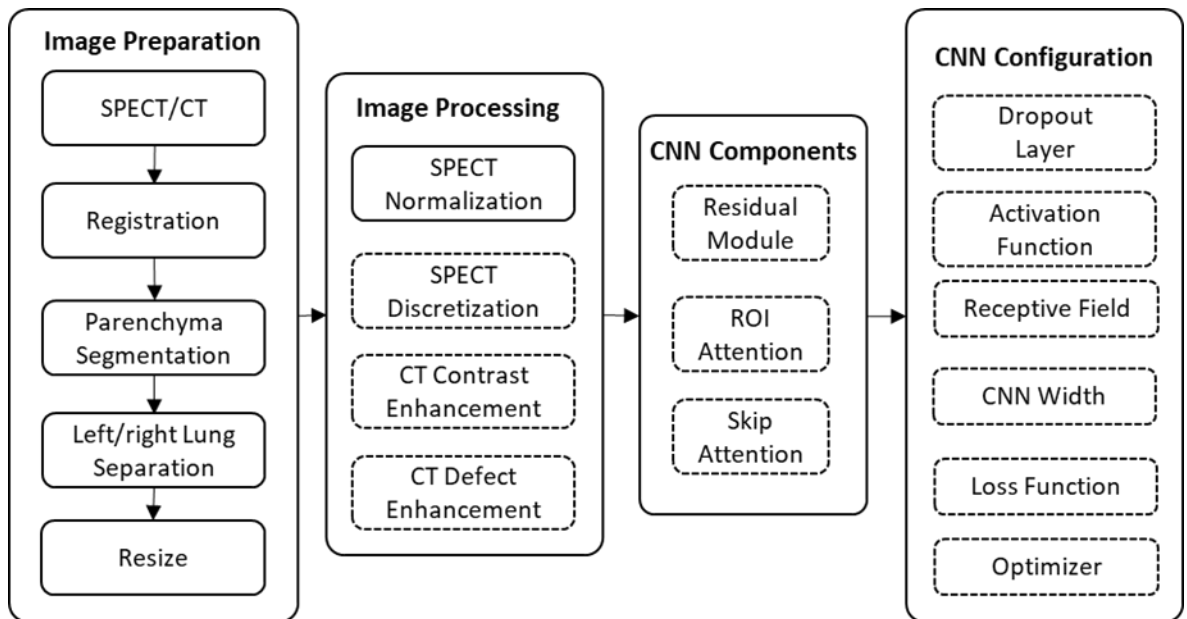
Figure 18 The workflow of this framework. The effects of the components in the

dashed outlined box were evaluated using ablation experiments.

3.4 *CTFM for Pulmonary CT-to-perfusion Translation*

3.4.1 Dataset and Study Design

The SPECT/CT images of 75 patients in the QMH perfusion dataset were used for this study. Two patients were excluded due to incomplete lung images. Among the 73 patients, 12 patients were diagnosed with lung cancer and 4 patients were diagnosed with no disease. The remaining patients had different types of lung diseases such as pulmonary hypertension, chronic pulmonary embolism, and pulmonary arteriovenous malformation.

The effectiveness of this framework was evaluated using a training–validation–testing scheme (Figure 19). To ensure the reliability of the test performance in a small dataset, a relatively large percentage of the test set (~40%) was randomly selected from the entire patient cohort. The remaining patients were used as the training set (~60%), and were randomly grouped into four folds for cross validation. In each validation, we used three folds of training samples to derive the network weights, and the remaining fold for validation. The cross-validation results were used to tune the hyperparameters of the model. It was then fitted to the entire training set to afford the final model, and its performance was evaluated on the test set.
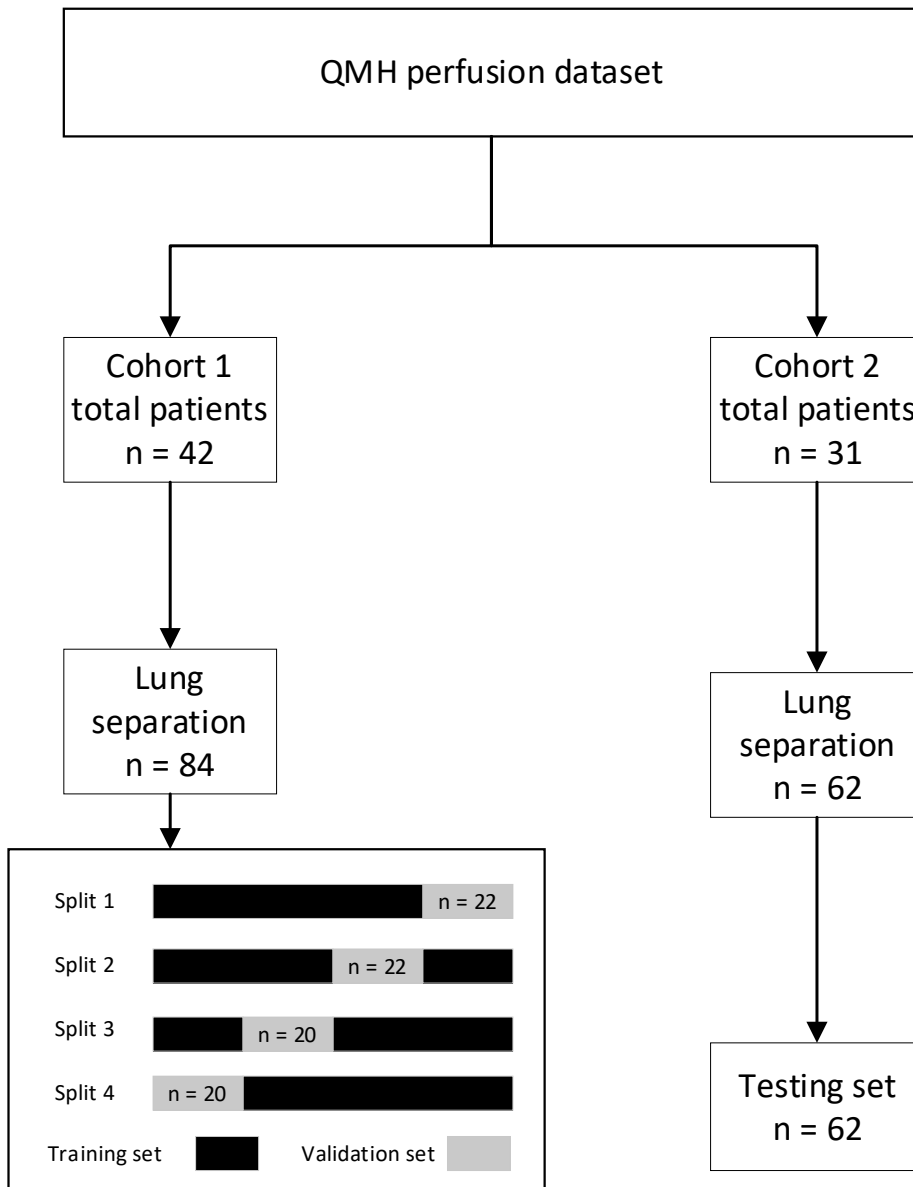
Figure 19 Study design for the pulmonary CT-to-perfusion Translation

3.4.2 Evaluation Metrics

Voxel-wise Agreement Function-wise Concordance: The performance of the DL-CTFM method was evaluated using Spearman's correlation coefficient (R) for intensity correlation assessment (Equation 4) and structural similarity index (SSIM) for perceptual similarity assessment (Equation 5). The generated DL-CTFM images were up-sampled to the original SPECT resolution and compared with the normalized SPECT perfusion.

Function-wise Concordance: The function-wise concordance between the ground-truth SPECT images and DL-CTFM generated images was evaluated to validate its potential applicability of DL-CTFM in FLART. The generated SPECT perfusion images was compared with the normalized SPECT perfusion images directly. The threshold to separate the high functional lung from the low lung functional lung was set to 0.66. Similar approaches have been used in a lung ventilation study [108] and suggested for FLART treatment plan optimization [56]. Subsequently, we evaluated the dice similarity coefficient (*DSC*) between the low/high functional volumes picked out by threshold value from the predict and ground-truth perfusion images (Equation 6):

$$DSC = \frac{2 * |p \cap y|}{|p| + |y|}, \qquad\qquad 6$$

where p is functional/defect volume in the predicted perfusion images, and y is the volume in the ground-truth SPECT perfusion images. The overall concordance can be inferred as the mean value of *DSC* among all the evaluated testing cases.

3.4.3 Implementation

Each layer is updated using error backpropagation with an adaptive moment estimation optimizer (ADAM). The epoch number is set to 500, such that the loss function change in the pilot study was less than 0.2% for the subsequent five epochs. We implemented our network using the PyTorch 1.1 framework. All of the preprocessing steps were coded in Python. All of the experiments were performed using a workstation with central-processing Intel Core i7-8700 @ 3.2GHz, a graphics-processing unit NVIDIA GTX 2080 TI with 11 GB of memory and 32 GB of RAM.

3.5 Preliminary Study on Pulmonary CT-to-Ventilation Translation

Because of the COVID-19 pandemic, the ventilation scan has been suspended in the hospital. As a result, only 13 SPECT/CT ventilation images were collected for a preliminary evaluation of pulmonary CT-to-Ventilation translation. The patients in the QMH ventilation dataset had different types of lung diseases such as pulmonary hypertension, chronic pulmonary embolism, and pulmonary arteriovenous malformation. After the image preparation and image processing pipeline, 26 cases were randomly divided into a training group (n=12) and a testing group (n=14). However, the ventilation dataset is relatively small to train a deep learning model. To fully utilize features learned from the previous study, a transfer learning strategy was used to reuse information learned from previous CT-to-perfusion translation tasks for the new CT-ventilation translation task [109]. Two transfer learning methods were used in the ventilation study: (1) Fine-tune. The weighting of the ARNN layers learned from the perfusion dataset was adopting the new model. The entire ARNN model was unfrozen and re-training on the new ventilation datasets with a low learning rate. (2) Fixed extractor. After adopting the old weighting parameters, the layers of the encoder were fixed to avoid destroying any of the information. Then the layers of the decoder were trained on the new ventilation datasets. In the implementation, the early-stop method was used to avoid overfitting on the small ventilation dataset. The regularization technique of early stopping was used to avoid overfitting on the training set. The evaluation metrics and other settings are the same as the previous CT-to-perfusion translations study.

4.   Results

4.1   *Effects Analysis of DL-CTFM Framework*

4.1.1   Effects of Image Processing

The effects of image processing were firstly analyzed by removing different processing procedures using the proposed CNN model. Removing the CT contrast enhancement step, the overall performance of the framework was jeopardized by ~11%. Removing the median filter degraded the overall performance by ~5%. Removing the uniform filter affected the overall performance by ~2% reduction. Interestingly, removing all the four image processing components also led to an approximately 11% decline in the framework performance. Removing SPECT discretization decreased the correlation by 3% and increased the SSIM by ~4%, resulting in a ~ 5 ‰ increment of the overall performance.

To further explore the effects on images, two representing cases with a sharp defect and a gradient defect, respectively, were visualized for qualitative analysis (Figure 20). For the sharp defect case, predicted perfusions of all the scenarios have a relatively high functional regions (~0.7) on the upper lobe. The scenario of CTFM-med. generates the largest defect region in this group. For the gradient defect case, the scenario CTFM generates the largest defect region in this group. For both cases, removing the CT contrast enhancement causes an overestimation of the value in the defect region.
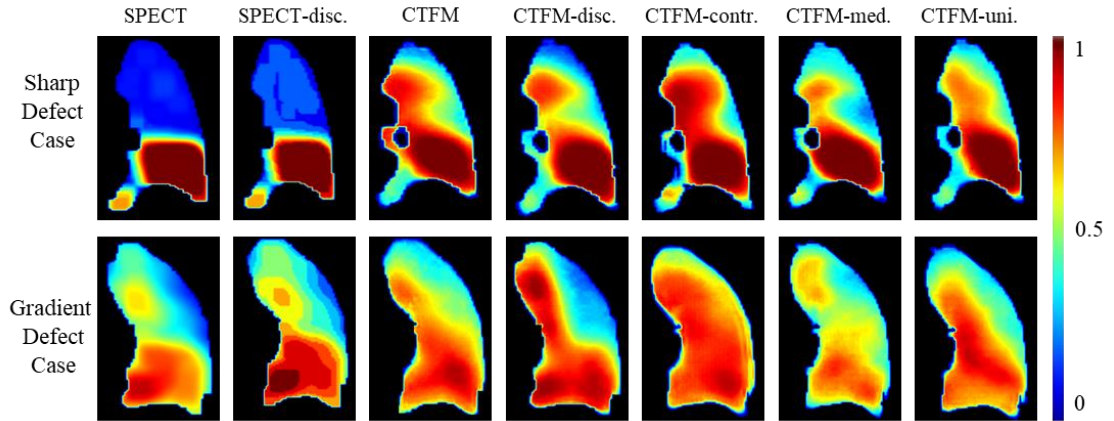
Figure 20 Illustration of the effects of different image processing steps in coronal view. Two representative cases with the sharp low functional and gradient low functional were visualized for comparison. The SPECT perfusion was normalized using the approach described in the method section. The color means the perfusion level. SPECT-disc. was the discretized label. CTFM was generated using the proposed setting. The following CTFMs were generated with five ablation scenarios: disc. – remove label discretization; contr. – remove CT contrast enhancement; med. – remove the median filter in CT low functional enhancement; uni. – remove the uniform filter in CT low functional enhancement.

Table 3 The effects of steps in the image processing pipeline

Ablation Experiments	*R*		*SSIM*		*SSIM+R*	Percent Difference
	Average	SD	Average	SD		
CTFM	0.6655	0.1351	0.7077	0.0740	1.3732	/
CTFM-disc.	0.6468	0.1697	0.7327	0.0899	1.3795	0.46%
CTFM-contr.	0.5702	0.2157	0.6520	0.0850	1.2222	-11.00%
CTFM-med.	0.6158	0.1869	0.6869	0.0798	1.3027	-5.13%
CTFM-uni.	0.6473	0.1535	0.6939	0.0662	1.3412	-2.33%

4.1.2 Effects of CNN Components

During the evaluation of the effects of the varying CNN components and configurations, the BCE loss was used. The proposed components and configurations, presented in bold in Table 4, yielded a performance score of 1.3703. By altering the maximum layer width of CNN, we can infer that optimal architecture has 128 filters (max CNN width 4×32) in the middle of the network. Increasing the kernel size from 3×3×3 to 5×5×5 resulted in a 3% improvement in performance. Using the dropout layer with a parameter of 0.1 increased the performance by 4.3% compared with the model without the dropout layer. Compared with ReLU or LeakyReLU (0.5), the performance with PReLU increased by 5.9% and 2.4%, respectively. The ROI attention, skip attention, and the residual module increased the performance by 4.4%, 3.1%, and 5.4%, respectively. In the visualization analysis, predicted perfusions of all CTFM scenarios are able to predict the defect region. Most scenarios generated a small relative high function region in the middle of the defect (Figure 21).
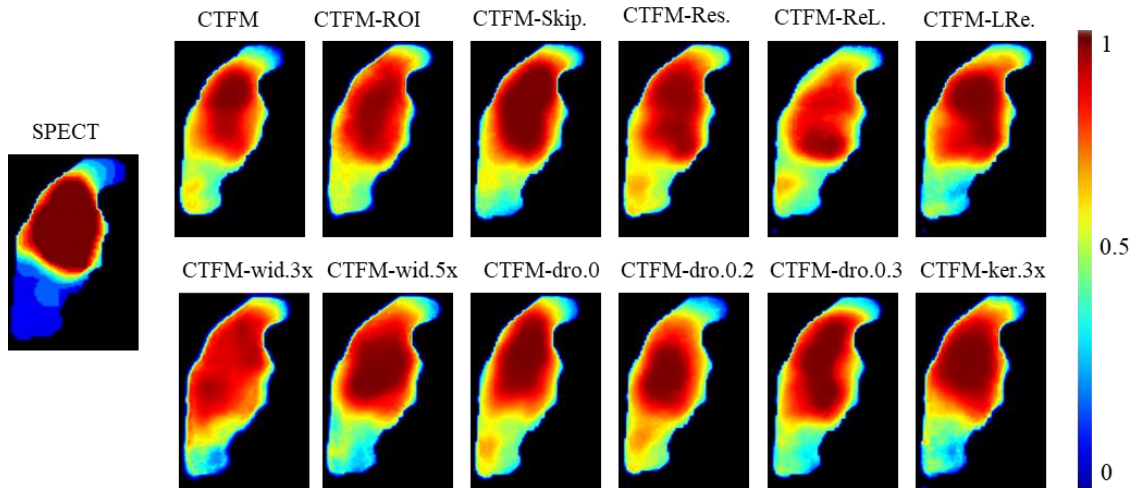
Figure 21 Illustration of the effect of varying CNN components and configurations of the representing case in coronal view. CTFM was generated using the proposed setting. The following CTFMs were generated in different CNN components and configurations: ROI. – remove ROI attention module, Skip. – remove skip attention module; Res. –remove residual module; ReL. – use ReLU instead of PReLU; LRe. – use LReLU instead of PReLU; wid. – with CNN width of 3, 4, or 5 times; dro – with dropout rate of 0, 0.2, or 0.3; ker 3x– use kernel size of 3×3×3 install of 5×5×5.

Table 4 Performance of the proposed network with different CNN components and

configurations

CNN Width	Kernel Size	Dropout Rate	ROI Attention	Skip Attention	Residual Module	Activation Function	*SSIM*	*R*	*SSIM+R*	Percent Difference
3×32	5×5×5	0.1	1	1	1	PReLU	0.6757	0.6305	1.3062	-4.68
4×32	**5×5×5**	**0.1**	**1**	**1**	**1**	**PReLU**	**0.7052**	**0.6651**	**1.3703**	**0**
5×32	5×5×5	0.1	1	1	1	PReLU	0.6831	0.6272	1.3103	-4.38
4×32	5×5×5	0	1	1	1	PReLU	0.6884	0.6236	1.3120	-4.25
4×32	5×5×5	0.2	1	1	1	PReLU	0.7014	0.6634	1.3648	-0.40
4×32	5×5×5	0.3	1	1	1	PReLU	0.7006	0.6519	1.3525	-1.30
4×32	3×3×3	0.1	1	1	1	PReLU	0.6867	0.6361	1.3228	-3.47
4×32	5×5×5	0.1	1	1	1	LReLU	0.6911	0.6458	1.3369	-2.44
4×32	5×5×5	0.1	1	1	1	ReLU	0.6882	0.6012	1.2894	-5.90
4×32	5×5×5	0.1	/	1	1	PReLU	0.6819	0.6276	1.3095	-4.44
4×32	5×5×5	0.1	1	/	1	PReLU	0.6879	0.6400	1.3279	-3.09
4×32	5×5×5	0.1	1	1	/	PReLU	0.6877	0.6085	1.2962	-5.41

4.1.3 Effects of Loss Functions and Optimizers

Table 5 demonstrates the effects of using different optimizers and loss functions for network training. Using Adam optimizer to minimize the BCE loss function yielded the best performance score. Figure 22 illustrates the convergence of different optimizers in terms of the validation loss over the number of epochs. Adam resulted in the fastest initial descent, followed by AdaGrad. We also tested the performance of auxiliary loss addition. The network performance increased by 2‰ by adding auxiliary correlation to BCE loss and 1% to MSE loss. Adding SSIM auxiliary loss to BCE decreased the correlation index by 10% and the overall performance by 7%. For the MSE loss, adding the SSIM auxiliary loss dampened the performance by 1‰.

Table 5 The effects of using different optimizers and loss functions

Optimizer	Loss Function	*SSIM*	*R*	*SSIM* + *R*
SGD	BCE	0.6518	0.5897	1.2415
AdaGrad	BCE	0.6168	0.6496	1.2664
Adam	BCE	0.7052	0.6651	1.3703
AdaGrad	MSE	0.6152	0.6315	1.2467
SGD	MSE	0.6519	0.5782	1.2301
Adam	MSE	0.7014	0.6408	1.3422
Adam	**BCE + *R***	**0.7077**	**0.6655**	**1.3732**
Adam	MSE + *R*	0.6999	0.6560	1.3559
Adam	BCE+SSIM	0.6668	0.5960	1.2628
Adam	MSE+SSIM	0.6941	0.6463	1.3404
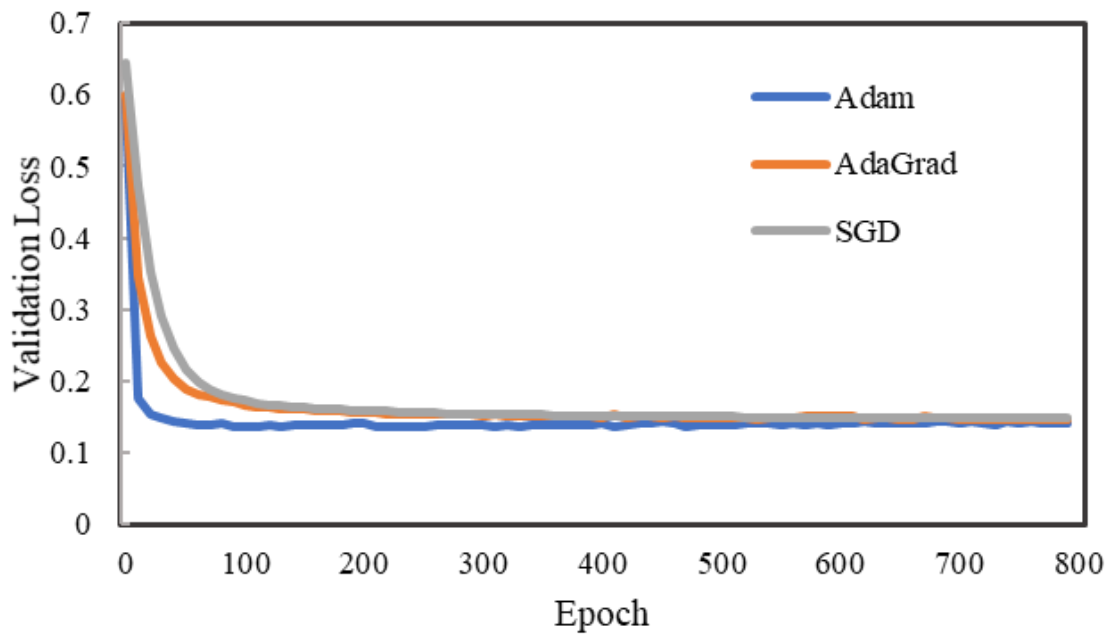
Figure 22 Comparison of the convergence speed with different optimizers.

4.1.4 Overall Effects Analysis

The performance of the final CTFM framework was illustrated in Figure 23. For the testing group, the proposed CTFM framework achieved an average DSC value of 0.8120 ± 0.0789 for high functional lung, average DSC value of 0.6682 ± 0.0867 for low functional lung, the average R value of 0.6534 ± 0.1432, and average SSIM value of 0.7437 ± 0.0739. The normalized SPECT perfusion and CTFM perfusion of representative cases in the testing group were visualized in coronal views in Figure 24 for qualitative evaluations. In the 22 testing cases, 10 cases (45.5%) have correlation values larger than 0.7; 6 cases (27.3%) have correlation values between 0.6 to 0.7; 6 cases (27.3%) have correlation values less than 0.6.

Compared with the widely used U-Net using identical image processing procedures (Figure 23), the proposed perfusion mapping CNN model outperformed U-Net by 0.5 % (average DSC value of high functional lung, $p = 0.3077$), 4.8% (average DSC value of low functional lung, $p = 0.0188$), 6.3% (average R value, $p = 0.0423$), and 3.0% (average SSIM value, $p = 0.0002$). The proposed CNN model can significantly reduce the computational time (2.5h) compared with U-Net (11.3h). Additionally, removing the image processing procedures in the framework decreased the four metrics by 2.8%, 12.4%, 12.2%, and 4.2%, respectively.
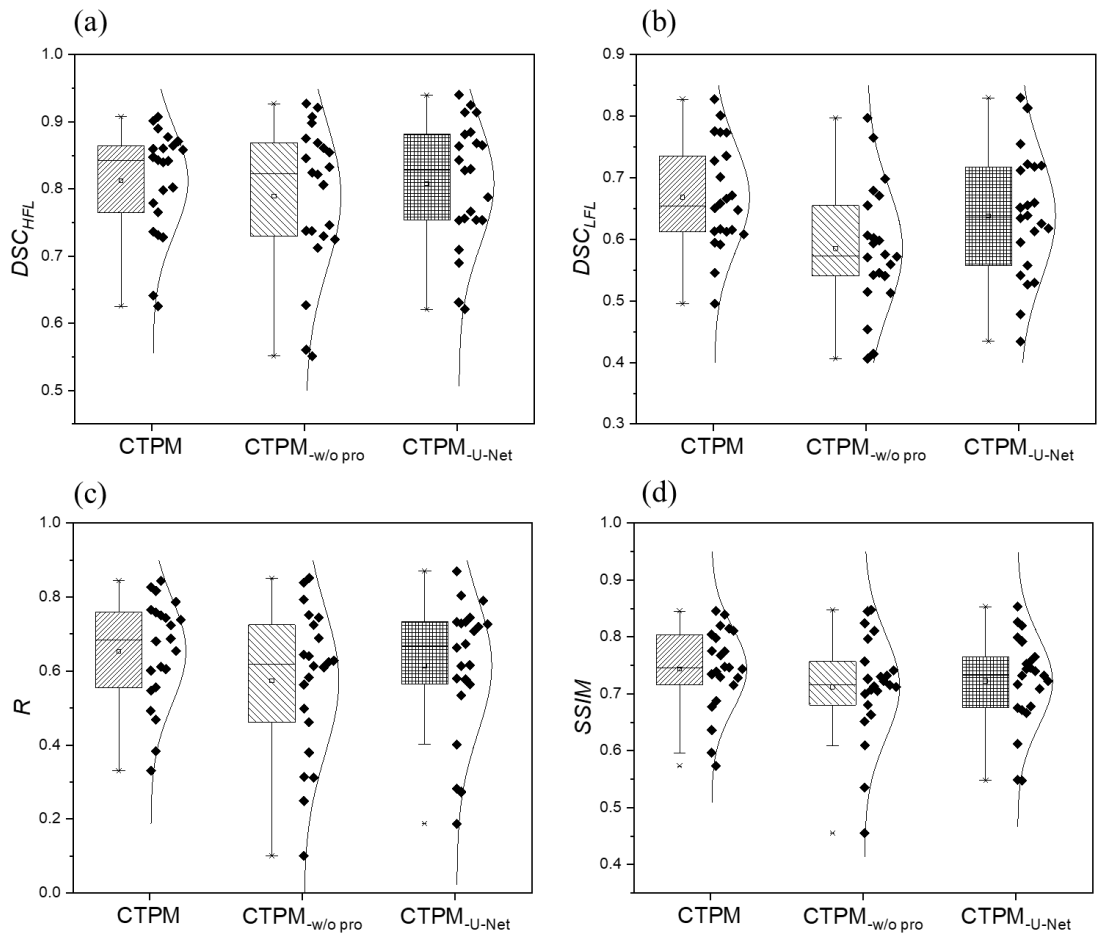
Figure 23 Overall performance analysis of the CTFM framework for the testing

group. (a) DSC of the high functional lung. (b) DSC of the low functional lung. (c)

Correlation coefficient. (d)Structural similarity. CTFM$_{-w/o\ pro}$ indicates prediction

without image processing. CTFM$_{-U-Net}$ indicates prediction using U-Net.
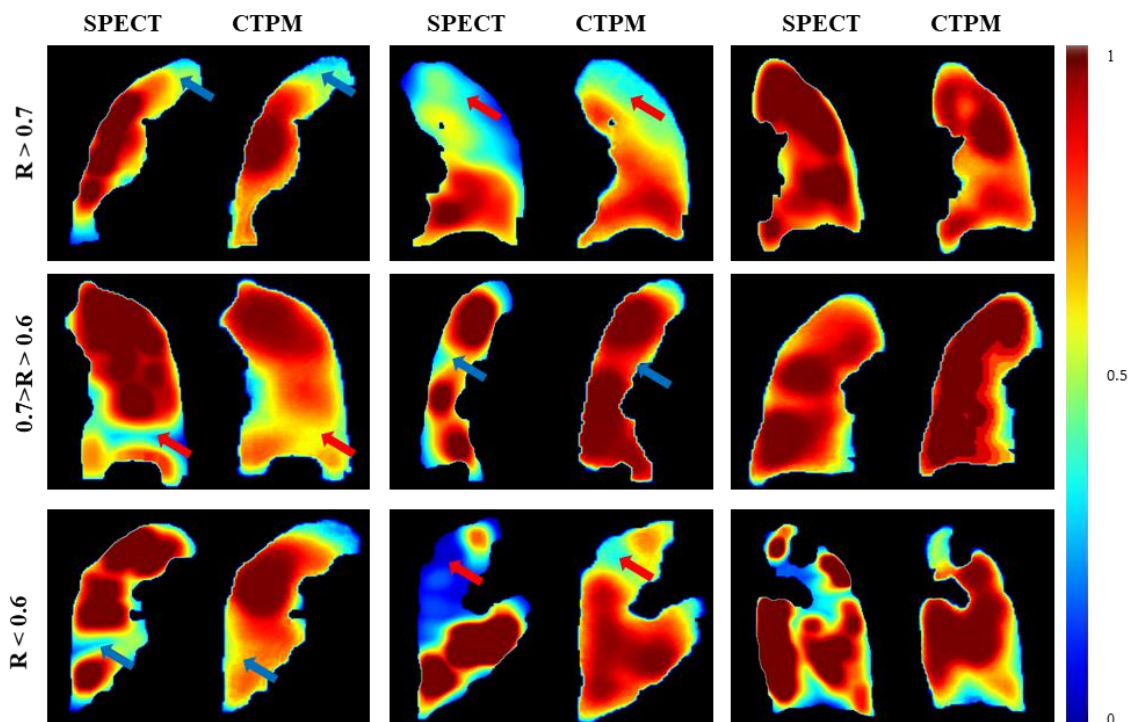
Figure 24 Comparison of the SPECT perfusion and CTFM perfusion of representative cases in the testing group. The arrows point to the primary low functional regions.

4.2   *Pulmonary CT-to-Perfusion Translation*

4.2.1   Cross Validation and Testing Results

Figure 4 demonstrated the evolution of the two evaluating metrics correlation coefficient R and structural similarity SSIM for voxel-wise agreement on the four validation folds and the testing set. For both R and SSIM metrics, the model performances on the testing group are within the variance range of the validation groups (Figure 25-a and -b), suggesting the high robustness of the proposed method.

The performances of the final model were visualized by the two box plots (Figures 4-c and -d). For the validation folds, the DL-CTFM perfusion images showed a moderate-to-high correlation (R = 0.6906 ± 0.1425) and high structural similarity (SSIM = 0.7704 ± 0.0703) with the corresponding reference perfusion images. For the testing set, the performance was similar to those in the validation folds, achieving R of 0.6733 ± 0.1728 and SSIM of 0.7635 ± 0.0697. For both R and SSIM, the unpaired two-tailed t-test showed no statistically significant difference (p = 0.2212 for R, p = 0.5037 for SSIM) and between validation and testing, suggesting the high generalizability of the trained model. In the testing set, only three outliers and one outlier (outside the 1.5 interquartile range) out of the 62 testing cases were observed for R and SSIM, respectively.
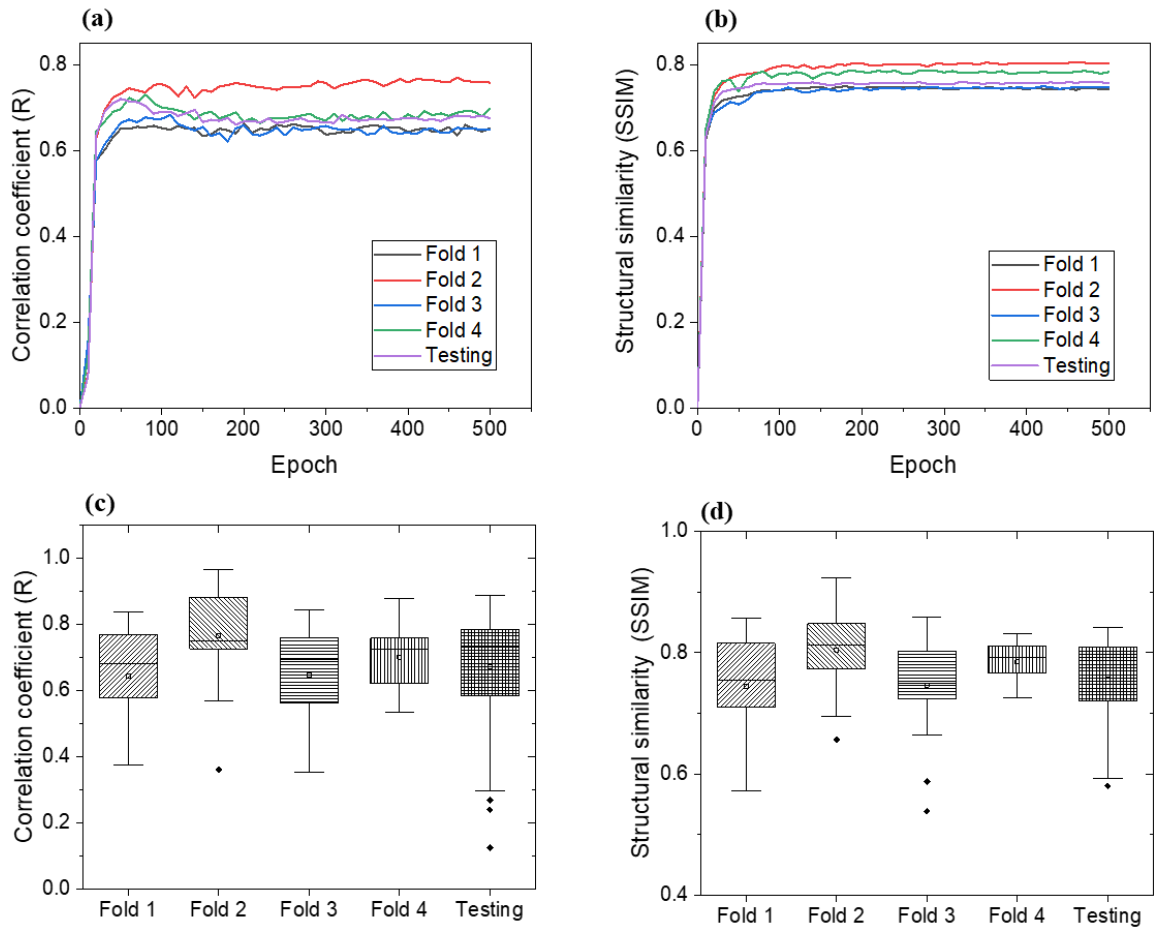
Figure 25 Results of pixel-wise agreement (correlation coefficient R and structural similarity index measure SSIM) evaluation for the 4 validation folds and the testing set. (a) and (b) are the evolutions of R and SSIM between the predicted and ground-truth perfusion images with increasing epochs during modeling. (c) and (d) are box plots of the pixel-wise agreement evaluation results of the final model.

4.2.2 Visualization Analysis

Figure 26 showed three representative testing cases, categorized as low-performance, average-performance, and high-performance. For the low-performance case, the low functional region in the middle of the lung was predicted as a high functional region in the corresponding DL-CTFM perfusion image. The low functional regions illustrated by the arrows were successfully predicted for the average-performance and high-performance cases.
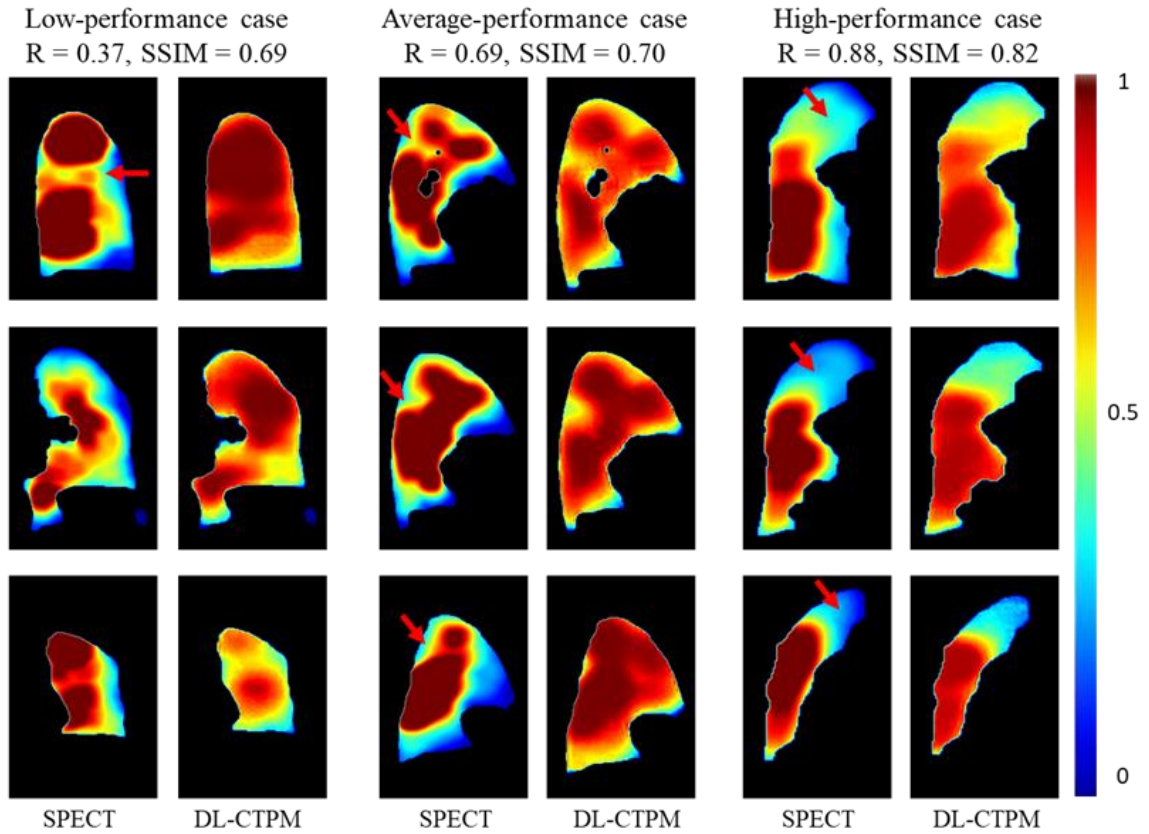
Figure 26 Comparison of SPECT perfusion and DL-CTFM perfusion of 3 representative cases. Each case is presented by 3 slices in either coronal or sagittal view. The red arrow indicates the main low functional region. The left 2 columns show a low-performing case, the middle 2 columns show an average case, and the right 2 columns show a high-performance case. The original SPECT images were normalized to the 75th percentile value without intensity discretization, as the predicted image has continuous pixel values.

4.2.3 Function-wise concordance

The DSC distributions for high and low functional lung volumes in the testing group are shown in Figure 27. In the 62 testing cases, 98% have a high functional lung DSC value larger than 0.6, and 71% have a low functional lung DSC value larger than 0.6. 94% of the testing patients showed high concordance (DSC > 0.7) between the high functional volumes contoured from the predicted and ground-truth perfusion, and 35% for the low functional volumes. The mean DSC value of high functional lung was 0.8183 ± 0.0752, ranging from 0.5819 to 0.9255; the mean DSC value of low functional lung was 0.6501 ± 0.1061, ranging from 0.2405 to 0.8212. Most testing cases showed high concordances for low/high functional volumes determined from the generated and ground-truth perfusion images.
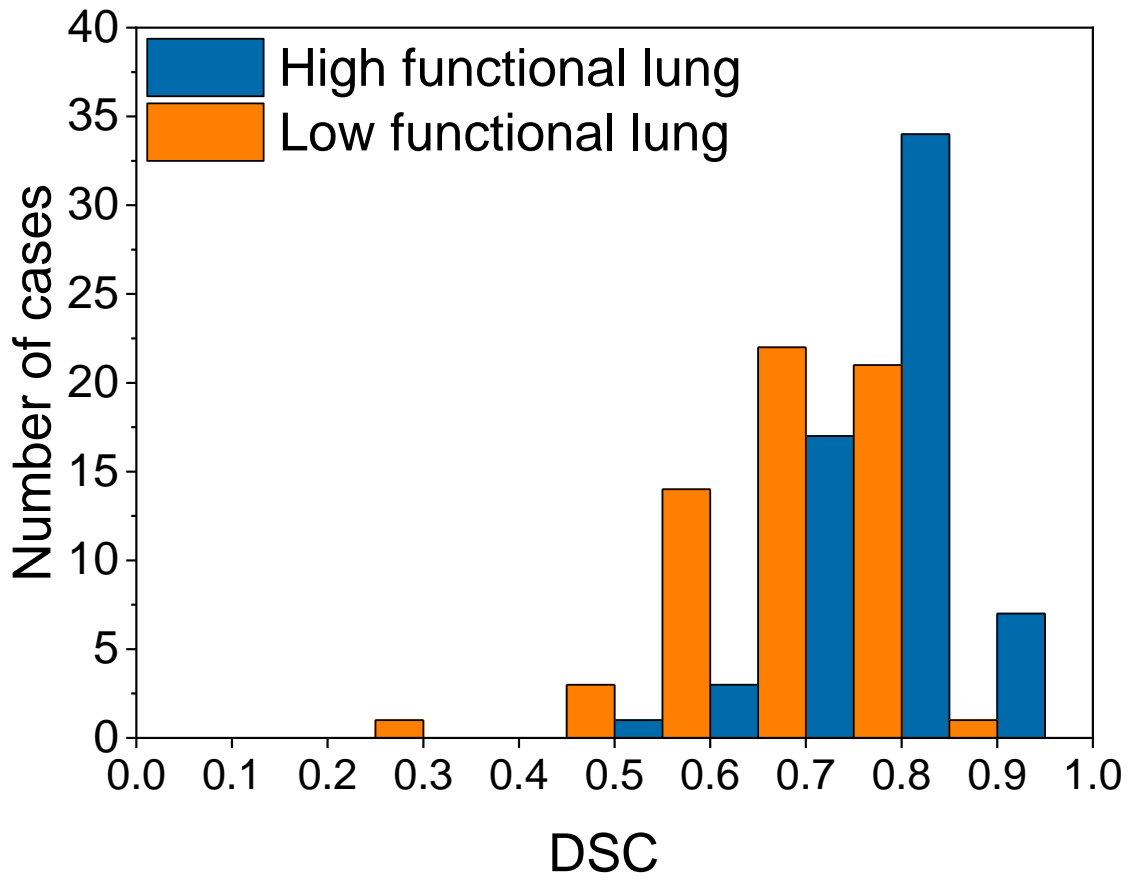
Figure 27 Histogram of function-wise concordance for the testing group.

4.3 *Preliminary Result of Pulmonary CT-to-Ventilation Translation*

4.3.1 Quantitative Evaluation

Figure 28 shows the evolution of the training and test loss with increasing training epochs. From epoch 452, the difference between the fitting curves of the training loss and test loss starts to decrease. Therefore, parameters at 452 epochs were chosen for the final model.
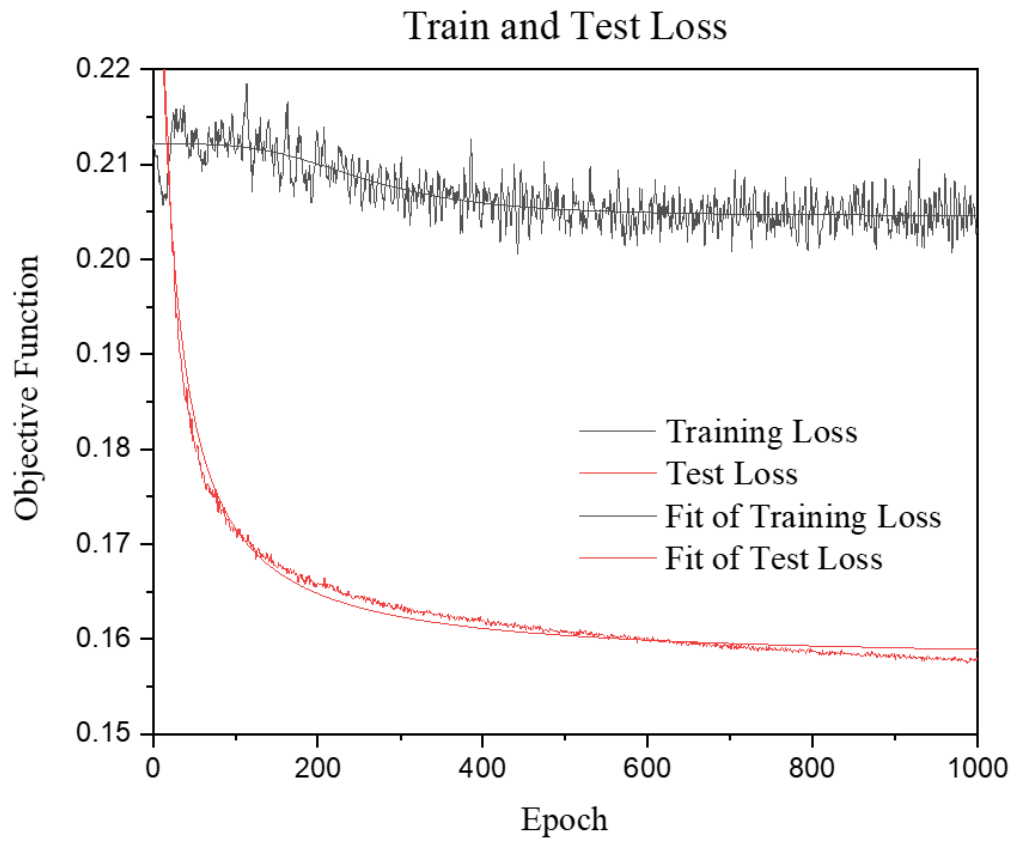
Figure 28 Training and test loss of pulmonary CT-to-ventilation translation

Figure 29 illustrates the box plots of the CTFM framework performance using four training strategies in three ventilation. The direct adoption of perfusion parameters has the least average performance in terms of correlation, function-wise concordance of low functional lung, and high functional lung (R= 0.4688 ± 0.1930; $DSC_{LFL} = 0.6850 \pm 0.1073$; $DSC_{HFL} = 0.6850 \pm 0.1073$). In terms of structural similarity, the fixed encoder model has the least average performance (SSIM = 0.6593 ± 0.0629). The fine-tune model has the best performance for all the four evaluation metrices (R= 0.6336 ± 0.0815; SSIM = 0.7391 ± 0.0681; $DSC_{LFL} = 0.7089 \pm 0.0438$; $DSC_{HFL} = 0.7265 \pm 0.0473$).
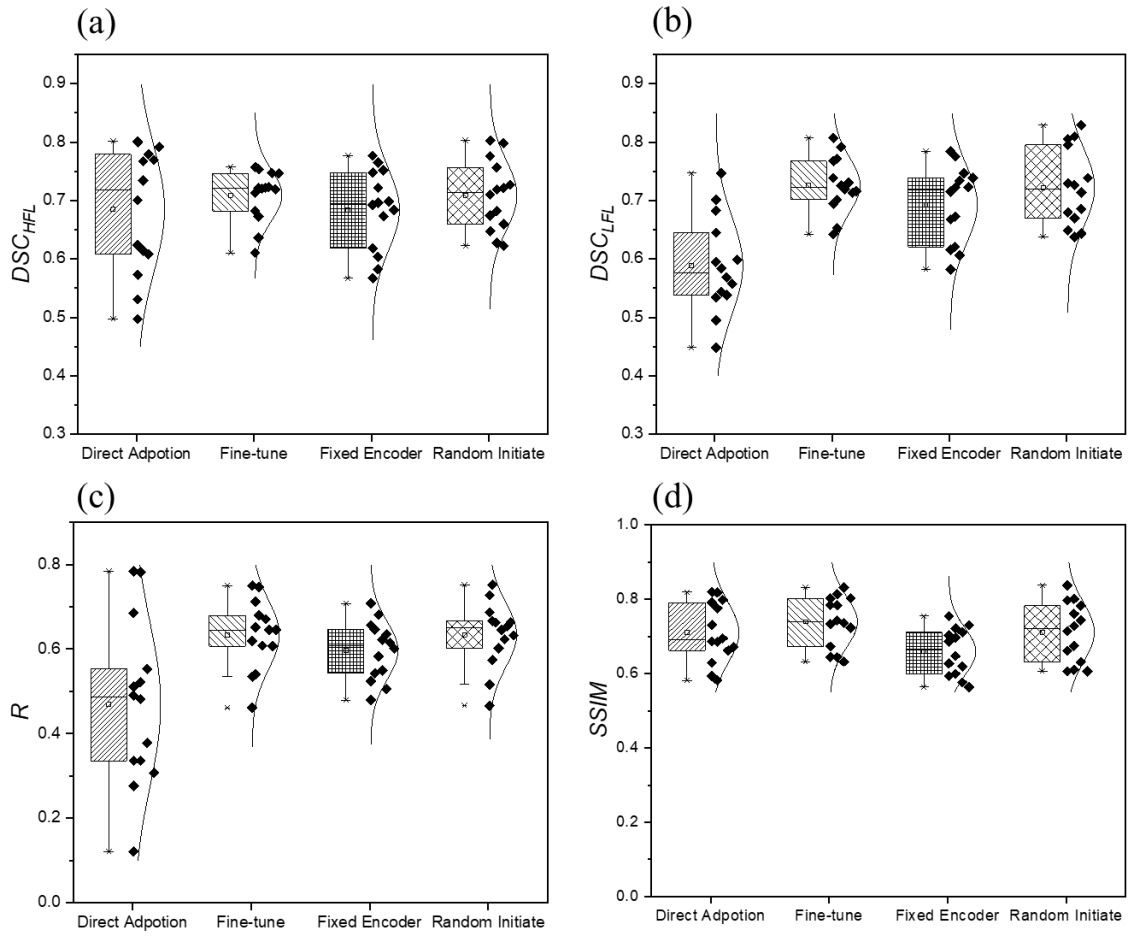
Figure 29 The Dice similarity coefficient (DSC) for high functional lung (HFL) and lung functional lung (LFL), correlation coefficient (R), and the structural similarity index measure (SSIM) of the CTFM framework for CT-to-ventilation translation using four training strategies.

4.3.2 Qualitative Evaluation

Figure 30 shows the impacts of different training strategies for the representative pulmonary hypertension case. For the tumor case, the low functional region in the upper lobe of the lung was predicted to be a high functional region in the corresponding direct

adoption perfusion image; in the other three training strategies (fine-tune, fixed extractor, and random initiate), the low-functional regions indicated by the arrow were successfully predicted. For the pulmonary hypertension case, the low-functional regions indicated by the arrow were successfully predicted in the fine-tuned model.
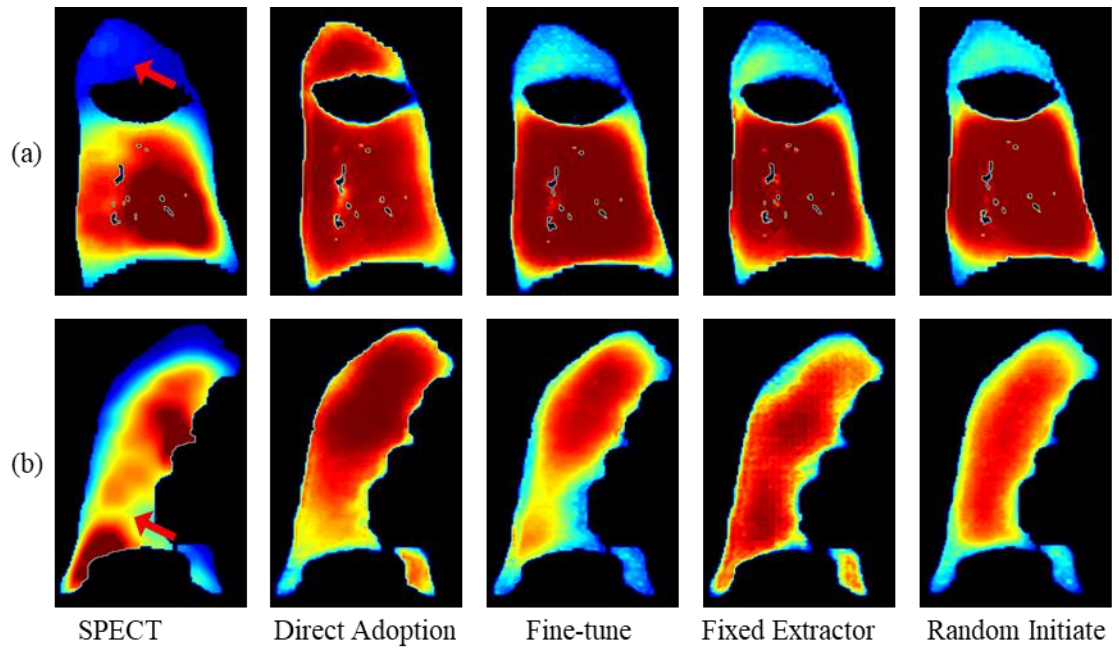
Figure 30 Impact of different training strategies for two representative cases. (a) case with lung tumor; (b) case with pulmonary hypertension. The red arrows indicate the main low functional regions. The dark oval-shaped region is the resected lung cancer region.

5.   Discussion

In this study, we constructed a deep learning-based CT function mapping (CTFM) framework for the pulmonary CT to functional map translation. This CTFM framework successfully generated perfusion-based lung functional images from the CT domain and demonstrated the feasibility of the ventilation-based lung functional images. This is the first study to perform CT-to-perfusion translation using a deep learning approach. A majority of the existing literature presented the CT-based function imaging in the form of ventilation. For instance, a multi-institution study on CT ventilation imaging showed a median correlation of 0.49 (range: 0.27-0.73) [57]. In our perfusion testing set, perfusion images generated using our DL-CTFM method achieved a moderate-to-high correlation (0.6733 ± 0.1728) to the ground-truth perfusion images. In terms of the functional-wise concordances, our findings demonstrated that 94% of testing cases had high functional lung *DSC* values of >0.7, and 71% had low functional lung *DSC* values of >0.6. This suggests that the regional functional lungs determined by the DL-CTFM method were in good agreement with those shown on the SPECT perfusion images. For the preliminary result in ventilation testing set, ventilation images generated using our CTFM framework average R of 0.6336 ± 0.0815, average SSIM of 0.7391 ± 0.0681, average DSC of 0.7089 ± 0.0438 for lung functional lung, average DSC of 0.7265 ± 0.0473 for high functional lung. Here, we note that both studies offer insightful implications to the field of radiotherapy, especially in view of evidence that both ventilation and perfusion imaging is correlated with the radiotherapy parameters, such as radiation pneumonitis [110]. Notably, results from this novel study is encouraging

91

and may serve as a stepping stone for achieving clinical applications in a broader scope using the synthesized lung functional images.

In our QMH perfusion dataset, the enrolled patients had a wide range of lung abnormalities (Table 1). The impact of this in the context of CT-to-perfusion translation also needs assessment. For instance, the original SPECT perfusion values of patients with pulmonary hypertension (ranging from 1,000 to 12,000) are much higher than those in patients with other lung conditions (typically less than 5,000). This abnormal distribution could diminish the learning ability of the ARNN model. To assess whether or not the model performance was degraded, for this reason, we performed an additional experiment excluding all patients with pulmonary hypertension in the dataset. Results showed a bit increment in R (by 4%) and negligible reduction in *SSIM* (by 1%), with corresponding values of 0.6983 ± 0.1392, and 0.7462 ± 0.071, respectively. This suggests that the heterogeneous distribution of lung abnormalities among patients may exert a mild impact on CT-to-perfusion translation, which may shed some light on the potential of our proposed method in diverse clinical diseases.

HU value is a function of the air/tissue ratio that determines the attenuation coefficient in the specified location. In the lung region, the fractional air/tissue of the parenchyma can also serve as a surrogate of the blood-gas exchange rate [65]. Therefore, pulmonary diseases that involve abnormal blood-gas exchange are generally manifested as textural alterations of the lung parenchyma. As shown in Figure 31, the low function region contoured on the SPECT perfusion cannot be observed on the greyscale CT image. After adjusting the window width and visualizing the image in a dense rainbow colormap, different patterns can be observed between the low and high

functional regions. Other studies have also shown that CT images can provide texture-based pulmonary function information [111] with high predictive power, which supports the feasibility of our CT-to-function translation. For example, the classification of interstitial lung diseases based on lung CT patterns achieved a high accuracy of 85.5% using a 6-layer neural network. These diseases are featured by inflammation and fibrosis of lung tissue. Meanwhile, typical disease patterns in CT are classified as reticulation, honey combing, ground glass opacity, consolidation, and micronodules [112].
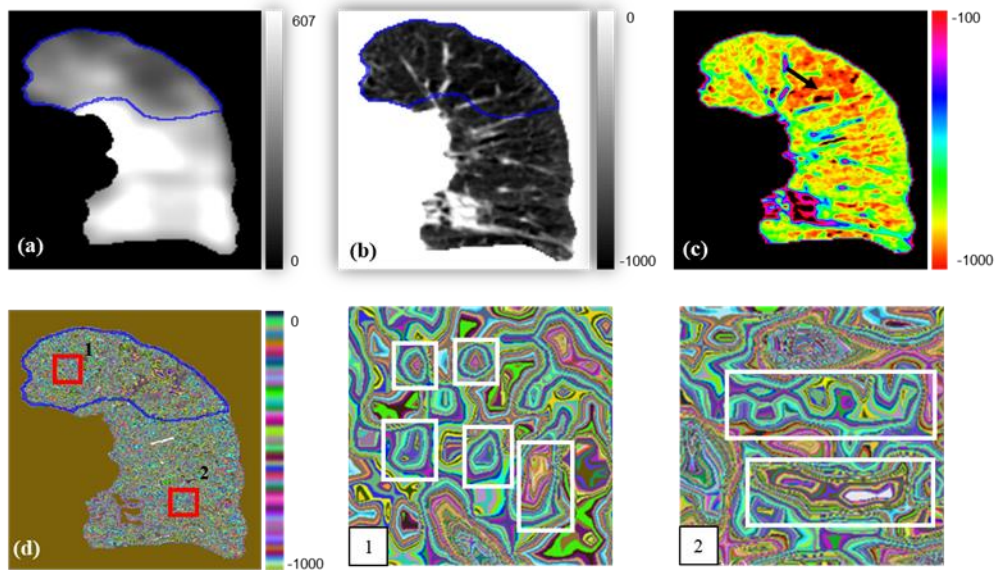
Figure 31 Illustration of healthy tissue and typical low function regions: (a) SPECT

perfusion image in a grey colormap. The blue contour indicates the clinically

significant low function region. (b) CT image in a grey colormap. (c) CT image in a

rainbow colormap. The black arrow indicates one defect region. (d) CT image in a

256 RGB colormap, which is a random selection of 256 RGB colors, used to

distinguish between a small number of labeled regions. Local magnification of (d)

indicates some typical texture patterns of low function region (1), and normal tissue

(2).

During the development of methodologies, improvements in results were observed while we kept optimizing the preprocessing procedures. For CT enhancement, histogram equalization and filtering were integrated and applied to CT images. The kernel sizes of both filters were adjusted according to the CNN architecture in the current study. The recommended values were 10 and 5 in ARNN for the median filter and uniform filter, respectively. For SPECT images, discretization and normalization procedures were applied. Discretization with 5 bins was used in a PET ventilation study [101]. To preserve useful information while removing noise signal during discretization process, 11 uniformly spaced bins were adopted in this study. In terms of normalization, different upper thresholds have been reported [65, 101-103]. In this study, a relatively small threshold in the high functional lung region is recommended for the CNN model due to the huge number of high voxel values in SPECT images, which may reduce the contribution of the features extracted from low functional regions to the loss function. Nevertheless, it should be noted that the determination of these preprocessing parameters should be dependent on the types of neural networks being trained. Further investigation of their impacts on model performance is warranted in the future, especially when a different neural network is adopted.

Noteworthy, the computational cost of 3D convolution neural networks (CNNs) and data scarcity are the commonly encountered challenges in deep learning study. 3D CNNs, such as GAN and GAN-based derivatives, are capable of fully exploiting the regional information in 3D images. However, it incurs significant computational demand due to the substantially increased amount of parameters, giving rise to a prolonged training time and reduced learning ability [105]. To overcome these

challenges, we designed an attention residual neural network (ARNN) with a reduced number of parameters. On the other hand, an inadequate amount of training samples incurs a higher risk of model overfitting, leading to false prediction. In this study, we separated the left and right lungs to augment the data size for model building. Besides, this separation also reduced the computational memory cost. In view of the anatomic difference between the left and right lungs, such as the number of lobes, we carried out a two-sample paired t-test to examine whether or not the anatomic difference between the left and right lungs affected CT-to-perfusion translation. Results from this analysis indicate that there was no statistically significant difference of both statistical estimates between the left/right lungs ($p = 0.876$ for R, $p = 0.4047$ for $SSIM$). This suggests that the performance of the ARNN may not be influenced by the anatomic/structural difference since CNN mainly captures texture-based features [113-115], or shape-based features [116-118]. However, physiological correlations exist between the left and right lungs and are hard to be captured by CNN models. This correlation should be considered in future clinical applications, such as lung functional avoidance radiotherapy (FLART) for lung cancer patients.

Despite the encouraging findings shown above, there is several limitations in our study, as described below:

Firstly, predictive performance in terms of the voxel-wise agreement can be worse in low-performance cases, where the low functional region in the middle of the lung was predicted as a high functional region in the corresponding DL-CTFM perfusion image, as illustrated in Figure 5. These low-performance cases can be partly attributed to the low occurrence and small size of the low functioning regions,

decreasing the contribution to the loss function. Further investigations into these low-performance cases are warranted to enhance the overall accuracy and utility of the deep learning model.

Secondly, although SPECT is favored for the diagnosis and radiation treatment assessment [43, 119], it does not quantify the concentration of radioactivity in absolute units due to factors such as attenuation, scatter, and partial volume effects. This leads to uncertainties in reflecting the actual perfusion conditions, which increased the value heterogeneity among patients. On the other hand, perfusion condition is only one of the indicators for assessing lung function. Even our preliminary study on ventilation dataset showed it is feasible of using CTFM framework for a CT-to-ventilation translation. However, a validation on more ventilation patients is still needed. Low functional lung regions predicted by both modalities should be integrated in treatment planning. Future studies translating CT images into both pulmonary perfusion and ventilation information can potentially enhance the value of the deep learning model in assessing the overall lung function.

Thirdly, the potential applicability of DL-CTFM in FLART was only evaluated by the structural concordance to the real SPECT images. Direct contextualization of the synthesized SPECT to the application of FLART requires validations in both dosimetry and clinical outcomes. Nevertheless, our findings highlighted the potential of CT-to-perfusion translation and may serve as a valuable steppingstone for achieving clinical application of FLART without additional functional imaging. Upon the foundation of our results, we encourage scientists in the field to carry out a large-cohort study specifically on lung cancer patients and incorporate information of dosimetry and

radiation-induced toxicities, whenever possible, to fully contextualize the results in the context of FLART application.

Lastly, as similar to other deep learning studies, the generalizability of our model can be influenced by heterogeneities in scanners (both CT and SPECT), imaging protocols (such as image acquisition and reconstruction settings), and image preprocessing steps (such as normalization and discretization). To achieve broader model deployment in future, a standardized protocol and/or a multi-institutional study is warranted to ensure model generalizability across various clinical settings.

6.   Conclusion

In this study, we developed a novel DL-CTFM framework for estimating lung functional images from the CT domain using a 3D ARNN. The deep convolutional neural network, in conjunction with an image processing pipeline for feature enhancement, is capable of feature extraction from CT images for perfusion synthesis. In this proposed framework, image processing, especially CT contrast enhancement, plays a key role in the synthesis. The modules and configurations of the CNN model (residual module, ROI attention, skip attention, dropout out, activation function) were also important to improve the prediction performance. This CTFM framework provides insights for relevant research studies in the future and enables other researchers to leverage for the development of optimized CNN models for CTFM. In the CT-to-perfusion translation, the CTFM framework yields moderate-to-high voxel-wise approximations of lung perfusion. To further contextualize these results toward future clinical application, a multi-institutional large-cohort study is warranted. In the CT-to-ventilation translation, the preliminary results demonstrate the feasibility of CT-to-ventilation translation. Compared to current methods of pulmonary function imaging, our method is easier, faster, less costly, and more accessible for clinically implementation of functionally guided radiotherapy. Owning to these unique advantages, our method holds great promises to expedite and expand the clinical adaption of functionally guided radiotherapy.

7. Reference

1.   Spurkland, A., *Gray's anatomy for students.* Tidsskrift for Den Norske Laegeforening, 2016. **136**(5): p. 456-456.

2.   Faiz M. Khan PhD, J.P.G.P., *The physics of radiation therapy.*

3.   Voicu-Maceseanu, A., et al., *[Epidemiology of lung cancer].* Pneumologia, 2007. **56**(2): p. 78-84.

4.   Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.* Int J Cancer, 2015. **136**(5): p. E359-86.

5.   Ferlay, J., et al., *Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods.* Int J Cancer, 2019. **144**(8): p. 1941-1953.

6.   Feng, R.M., et al., *Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics?* Cancer Commun (Lond), 2019. **39**(1): p. 22.

7.   Collins, L.G., et al., *Lung cancer: diagnosis and management.* Am Fam Physician, 2007. **75**(1): p. 56-63.

8.   Yoon, S.M., T. Shaikh, and M. Hallman, *Therapeutic management options for stage III non-small cell lung cancer.* World J Clin Oncol, 2017. **8**(1): p. 1-20.

9.   Vansteenkiste, J., et al., *Randomized controlled trial of resection versus radiotherapy after induction chemotherapy in stage IIIA-N2 non-small cell lung cancer.* Journal of Thoracic Oncology, 2007. **2**(8): p. 684-685.

10.   Albain, K.S., et al., *Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial.* Lancet, 2009. **374**(9687): p. 379-386.

11.   Laine, A.M., K.D. Westover, and H. Choy, *Radiation Therapy as a Backbone of Treatment of Locally Advanced Non-Small Cell Lung Cancer.* Seminars in Oncology, 2014. **41**(1): p. 57-68.

12.   Giuranno, L., et al., *Radiation-Induced Lung Injury (RILI).* Frontiers in Oncology, 2019. **9**.

13.   Mehta, V., *Radiation pneumonitis and pulmonary fibrosis in non-small-cell lung cancer: Pulmonary function, prediction, and prevention.* International Journal of Radiation Oncology Biology Physics, 2005. **63**(1): p. 5-24.

14.   Jain, V. and A.T. Berman, *Radiation Pneumonitis: Old Problem, New Tricks.* Cancers, 2018. **10**(7).

15.   Robnett, T.J., et al., *Factors predicting severe radiation pneumonitis in patients receiving definitive chemoradiation for lung cancer.* Int J Radiat Oncol Biol Phys, 2000. **48**(1): p. 89-94.

16.   Midha, A., S. Dearden, and R. McCormack, *EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII).* Am J Cancer Res, 2015. **5**(9): p. 2892-911.

17.   Palma, D.A., et al., *Predicting radiation pneumonitis after chemoradiation therapy for lung cancer: an international individual patient data meta-analysis.* Int J Radiat Oncol Biol Phys, 2013. **85**(2): p. 444-50.

18.   Hoover, D.A., et al., *Functional lung avoidance for individualized radiotherapy (FLAIR): study protocol for a randomized, double-blind clinical trial.* BMC Cancer, 2014. **14**: p. 934.

19.   Movsas, B., et al., *Quality of Life (QOL) Analysis of the Randomized Radiation (RT) Dose-Escalation NSCLC Trial (RTOG 0617): The Rest of the Story.* International Journal of Radiation Oncology Biology Physics, 2013. **87**(2): p. S1-S2.

20.   Zhou, Z.Y., et al., *Pulmonary emphysema is a risk factor for radiation pneumonitis in NSCLC patients with squamous cell carcinoma after thoracic radiation therapy.* Scientific Reports, 2017. **7**.

21.   Dieterich, S., et al., *Chapter 13 - Simulation for Radiotherapy Treatment Planning*, in *Practical Radiation Oncology Physics*, S. Dieterich, et al., Editors. 2016, Elsevier: Philadelphia. p. 180-188.

22.   Christian, J.A., et al., *The incorporation of SPECT functional lung imaging into inverse radiotherapy planning for non-small cell lung cancer.* Radiotherapy and Oncology, 2005. **77**(3): p. 271-277.

23.   Larici, A.R., et al., *Lung Abnormalities at Multimodality Imaging after Radiation Therapy for Non-Small Cell Lung Cancer.* Radiographics, 2011. **31**(3): p. 771-789.

24. Matuszak, M.M., et al., *Priority-driven plan optimization in locally advanced lung patients based on perfusion SPECT imaging.* Adv Radiat Oncol, 2016. **1**(4): p. 281-289.

25. Faught, A.M., et al., *Evaluating Which Dose-Function Metrics Are Most Critical for Functional Guided Radiation Therapy with CT Ventilation Imaging.* International Journal of Radiation Oncology Biology Physics, 2017. **99**(2): p. E454-E455.

26. Waxweiler, T., et al., *A complete 4DCT-ventilation functional avoidance virtual trial: Developing strategies for prospective clinical trials.* J Appl Clin Med Phys, 2017. **18**(3): p. 144-152.

27. Yamamoto, T., et al., *Changes in Regional Ventilation During Treatment and Dosimetric Advantages of CT Ventilation Image Guided Radiation Therapy for Locally Advanced Lung Cancer.* Int J Radiat Oncol Biol Phys, 2018. **102**(4): p. 1366-1373.

28. Powers, K.A. and A.S. Dhamoon, *Physiology, Pulmonary Ventilation and Perfusion*, in *StatPearls*. 2020: Treasure Island (FL).

29. Petersson, J. and R.W. Glenny, *Gas exchange and ventilation-perfusion relationships in the lung.* European Respiratory Journal, 2014. **44**(4): p. 1023-1041.

30. Behrman, R.E., R. Kliegman, and H.B. Jenson, *Nelson textbook of pediatrics*. 17th ed. 2004, Philadelphia, PA: Saunders. xlviii, 2618 p.

31. Nilsson, K.R., J. Piccini, and Johns Hopkins Hospital. Osler Medical Service., *The Osler medical handbook / the Osler Medical Service, The Johns Hopkins Hospital*

; *editors, Kent R. Nilsson, Jr., Jonathan Piccini*. 2nd ed. 2006, Philadelphia: Saunders/Elsevier. xxviii, 1039 p., 12 p. of plates.

32.   Huet, Y., et al., *Hypoxemia in Acute Pulmonary-Embolism.* Chest, 1985. **88**(6): p. 829-836.

33.   Enright, P., *Office-based DLCO tests help pulmonologists to make important clinical decisions.* Respiratory Investigation, 2016. **54**(5): p. 305-311.

34.   *Pulmonary circulation*. 2016, the Institute for Quality and Efficiency in Health Care: informedhealth.

35.   Lee, E., et al., *Functional lung avoidance and response-adaptive escalation (FLARE) RT: Multimodality plan dosimetry of a precision radiation oncology strategy.* Med Phys, 2017. **44**(7): p. 3418-3429.

36.   Kida, S., et al., *CT ventilation functional image-based IMRT treatment plans are comparable to SPECT ventilation functional image-based plans.* Radiother Oncol, 2016. **118**(3): p. 521-7.

37.   Yamamoto, T., et al., *Investigation of four-dimensional computed tomography-based pulmonary ventilation imaging in patients with emphysematous lung regions.* Phys Med Biol, 2011. **56**(7): p. 2279-98.

38.   Isawa, T., et al., *Technegas for inhalation lung imaging.* Nucl Med Commun, 1991. **12**(1): p. 47-55.

39.   Zhang, J.A., et al., *Radiation-Induced Reductions in Regional Lung Perfusion: 0.1-12 Year Data from a Prospective Clinical Study.* International Journal of Radiation Oncology Biology Physics, 2010. **76**(2): p. 425-432.

40.   Seppenwoolde, Y., et al., *Optimizing radiation treatment plans for lung cancer using lung perfusion information.* Radiotherapy and Oncology, 2002. **63**(2): p. 165-177.

41.   Mistry, N.N., et al., *Pulmonary perfusion imaging in the rodent lung using dynamic contrast-enhanced MRI.* Magnetic Resonance in Medicine, 2008. **59**(2): p. 289-297.

42.   Ascha, M., R.D. Renapurkar, and A.R. Tonelli, *A review of imaging modalities in pulmonary hypertension.* Annals of Thoracic Medicine, 2017. **12**(2): p. 61-73.

43.   Gefter, W.B. and H. Hatabu, *Functional lung imaging: emerging methods to visualize regional pulmonary physiology.* Acad Radiol, 2003. **10**(10): p. 1085-9.

44.   Robbins, M.E., et al., *Imaging radiation-induced normal tissue injury.* Radiat Res, 2012. **177**(4): p. 449-66.

45.   Yuan, S.G., et al., *Changes in Global Function and Regional Ventilation and Perfusion on SPECT During the Course of Radiotherapy in Patients With Non-Small-Cell Lung Cancer.* International Journal of Radiation Oncology Biology Physics, 2012. **82**(4): p. E631-E638.

46.   Scheenstra, A.E.H., et al., *Local dose-effect relations for lung perfusion post stereotactic body radiotherapy.* Radiotherapy and Oncology, 2013. **107**(3): p. 398-402.

47. Meng, X., et al., *Changes in Functional Lung Regions During the Course of Radiation Therapy and Their Potential Impact on Lung Dosimetry for Non-Small Cell Lung Cancer.* International Journal of Radiation Oncology Biology Physics, 2014. **89**(1): p. 145-151.

48. Bajc, M., et al., *EANM guideline for ventilation/perfusion single-photon emission computed tomography (SPECT) for diagnosis of pulmonary embolism and beyond.* European Journal of Nuclear Medicine and Molecular Imaging, 2019. **46**(12): p. 2429-2451.

49. Wheatley, J.R., *Spirometry: key to the diagnosis of respiratory disorders.* Med J Aust, 2017. **207**(10): p. 422-423.

50. Dubsky, S. and A. Fouras, *Imaging regional lung function: a critical tool for developing inhaled antimicrobial therapies.* Adv Drug Deliv Rev, 2015. **85**: p. 100-9.

51. Kolodziej, M., et al., *Lung function imaging methods in Cystic Fibrosis pulmonary disease.* Respir Res, 2017. **18**(1): p. 96.

52. Eslick, E.M., et al., *Measurement of preoperative lobar lung function with computed tomography ventilation imaging: progress towards rapid stratification of lung cancer lobectomy patients with abnormal lung function.* Eur J Cardiothorac Surg, 2016. **49**(4): p. 1075-82.

53.   Kipritidis, J., et al., *Validating and improving CT ventilation imaging by correlating with ventilation 4D-PET/CT using Ga-68-labeled nanoparticles.* Medical Physics, 2014. **41**(1).

54.   Le Roux, P.Y., et al., *PET/CT Lung Ventilation and Perfusion Scanning using Galligas and Gallium-68-MAA.* Seminars in Nuclear Medicine, 2019. **49**(1): p. 71-81.

55.   Tahir, B.A., et al., *Spatial Comparison of CT-Based Surrogates of Lung Ventilation With Hyperpolarized Helium-3 and Xenon-129 Gas MRI in Patients Undergoing Radiation Therapy.* Int J Radiat Oncol Biol Phys, 2018. **102**(4): p. 1276-1286.

56.   Yamamoto, T., et al., *Impact of four-dimensional computed tomography pulmonary ventilation imaging-based functional avoidance for lung cancer radiotherapy.* Int J Radiat Oncol Biol Phys, 2011. **79**(1): p. 279-88.

57.   Kipritidis, J., et al., *The VAMPIRE challenge: A multi-institutional validation study of CT ventilation imaging.* Med Phys, 2019. **46**(3): p. 1198-1217.

58.   Kuethe, D.O., et al., *Imaging lungs using inert fluorinated gases.* Magnetic Resonance in Medicine, 1998. **39**(1): p. 85-88.

59.   Edelman, R.R., et al., *Noninvasive assessment of regional ventilation in the human lung using oxygen-enhanced magnetic resonance imaging.* Nature Medicine, 1996. **2**(11): p. 1236-1239.

60.   Fain, S., et al., *Imaging of Lung Function Using Hyperpolarized Helium-3 Magnetic Resonance Imaging: Review of Current and Emerging Translational Methods and Applications.* Journal of Magnetic Resonance Imaging, 2010. **32**(6): p. 1398-1408.

61.   Mugler, J.P. and T.A. Altes, *Hyperpolarized 129Xe MRI of the human lung.* Journal of Magnetic Resonance Imaging, 2013. **37**(2): p. 313-331.

62.   Bailey, D.L., et al., *(68)Ga PET Ventilation and Perfusion Lung Imaging-Current Status and Future Challenges.* Semin Nucl Med, 2016. **46**(5): p. 428-35.

63.   Hofman, M.S., et al., *68Ga PET/CT ventilation-perfusion imaging for pulmonary embolism: a pilot study with comparison to conventional scintigraphy.* J Nucl Med, 2011. **52**(10): p. 1513-9.

64.   Guerrero, T., et al., *Dynamic ventilation imaging from four-dimensional computed tomography.* Phys Med Biol, 2006. **51**(4): p. 777-91.

65.   Kipritidis, J., et al., *Estimating lung ventilation directly from 4D CT Hounsfield unit values.* Med Phys, 2016. **43**(1): p. 33.

66.   Reinhardt, J.M., et al., *Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation.* Medical Image Analysis, 2008. **12**(6): p. 752-763.

67.   Simmons, A.B. and S.G. Chappell, *Artificial-Intelligence - Definition and Practice.* Ieee Journal of Oceanic Engineering, 1988. **13**(2): p. 14-42.

68.   Fazal, M.I., et al., *The past, present and future role of artificial intelligence in imaging.* European Journal of Radiology, 2018. **105**: p. 246-250.

69.   Kahn, C.E., *Artificial-Intelligence in Radiology - Decision-Support Systems.* Radiographics, 1994. **14**(4): p. 849-861.

70.   Dechter, R., *Learning While Searching in Constraint-Satisfaction-Problems*. 1986. 178-185.

71.   Aizenberg, I.N., *Neural networks based on multi-valued and universal binary neurons: Theory, application to image processing and recognition.* Computational Intelligence: Theory and Applications, 1999. **1625**: p. 306-316.

72.   Akkus, Z., et al., *Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions.* Journal of Digital Imaging, 2017. **30**(4): p. 449-459.

73.   Valueva, M.V., et al., *Application of the residue number system to reduce hardware costs of the convolutional neural network implementation.* Mathematics and Computers in Simulation, 2020. **177**: p. 232-243.

74.   Livne, M., et al., *A U-Net Deep Learning Framework for High Performance Vessel Segmentation in Patients With Cerebrovascular Disease.* Front Neurosci, 2019. **13**: p. 97.

75.   Milletari, F., N. Navab, and S.A. Ahmadi, *V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation.* Proceedings of 2016 Fourth International Conference on 3d Vision (3dv), 2016: p. 565-571.

76.   Krizhevsky, A., I. Sutskever, and G.E. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks.* Communications of the Acm, 2017. **60**(6): p. 84-90.

77.   Wang, T.H., et al., *A review on medical imaging synthesis using deep learning and its clinical applications.* Journal of Applied Clinical Medical Physics, 2021. **22**(1): p. 11-36.

78.   Han, X., *MR-based synthetic CT generation using a deep convolutional neural network method.* Medical Physics, 2017. **44**(4): p. 1408-1419.

79.   Jin, C.B., et al., *Deep CT to MR Synthesis Using Paired and Unpaired Data.* Sensors, 2019. **19**(10).

80.   Li, W., et al., *Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy.* Quantitative Imaging in Medicine and Surgery, 2020. **10**(6): p. 1223-1236.

81.   Dong, X., et al., *Synthetic MRI-aided multi-organ segmentation on male pelvic CT using cycle consistent deep attention network.* Radiotherapy and Oncology, 2019. **141**: p. 192-199.

82.   Stimpel, B., et al., *Projection Image-to-Image Translation in Hybrid X-ray/MR Imaging.* Medical Imaging 2019: Image Processing, 2019. **10949**.

83.   Zhong, Y., et al., *Technical Note: Deriving ventilation imaging from 4DCT by deep convolutional neural network.* Med Phys, 2019. **46**(5): p. 2323-2329.

84.   Ren, G., et al., *Deep Learning-Based Computed Tomography Perfusion Mapping (DL-CTPM) for Pulmonary CT-to-Perfusion Translation.* Int J Radiat Oncol Biol Phys, 2021.

85.   Qu, L.Q., et al., *Wavelet-based Semi-supervised Adversarial Learning for Synthesizing Realistic 7T from 3T MRI.* Medical Image Computing and Computer Assisted Intervention - Miccai 2019, Pt Iv, 2019. **11767**: p. 786-794.

86.   Liang, X., et al., *Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy.* Physics in Medicine and Biology, 2019. **64**(12).

87.   Nie, D., et al., *Medical Image Synthesis with Deep Convolutional Adversarial Networks.* IEEE Trans Biomed Eng, 2018. **65**(12): p. 2720-2730.

88.   Long, J., E. Shelhamer, and T. Darrell, *Fully Convolutional Networks for Semantic Segmentation.* 2015 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr), 2015: p. 3431-3440.

89.   Ronneberger, O., P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation.* Medical Image Computing and Computer-Assisted Intervention, Pt Iii, 2015. **9351**: p. 234-241.

90.   Çiçek, Ö., et al. *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation.* 2016. arXiv:1606.06650.

91.   Oktay, O., et al. *Attention U-Net: Learning Where to Look for the Pancreas.* arXiv e-prints, 2018. arXiv:1804.03999.

92.   Goodfellow, I.J., et al., *Generative Adversarial Nets.* Advances in Neural Information Processing Systems 27 (Nips 2014), 2014. **27**.

93.   Li, J., J.J. Jia, and D.L. Xu, *Unsupervised Representation Learning of Image-Based Plant Disease with Deep Convolutional Generative Adversarial Networks.* 2018 37th Chinese Control Conference (Ccc), 2018: p. 9159-9163.

94.   Chen, H., et al., *Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network.* IEEE Trans Med Imaging, 2017. **36**(12): p. 2524-2535.

95.   Wolterink, J.M., et al., *Generative Adversarial Networks for Noise Reduction in Low-Dose CT.* IEEE Trans Med Imaging, 2017. **36**(12): p. 2536-2545.

96.   Leung, K., et al., *A deep-learning-based fully automated segmentation approach to delineate tumors in FDG-PET images of patients with lung cancer.* Journal of Nuclear Medicine, 2018. **59**.

97.   Shin, H.C., et al., *Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning.* Ieee Transactions on Medical Imaging, 2016. **35**(5): p. 1285-1298.

98.   Kikinis, R., S.D. Pieper, and K.G. Vosburgh, *3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support*, in *Intraoperative Imaging and Image-Guided Therapy*, F.A. Jolesz, Editor. 2014, Springer New York: New York, NY. p. 277-289.

99.   Estepar, R.S., et al., *Chest Imaging Platform: An Open-Source Library And Workstation For Quantitative Chest Imaging.* American Journal of Respiratory and Critical Care Medicine, 2015. **191**.

100. Lehr, J.L. and P. Capek, *Histogram equalization of CT images.* Radiology, 1985. **154**(1): p. 163-9.

101. Kipritidis, J., et al., *Validating and improving CT ventilation imaging by correlating with ventilation 4D-PET/CT using 68Ga-labeled nanoparticles.* Med Phys, 2014. **41**(1): p. 011910.

102. Hegi-Johnson, F., et al., *Evaluating the accuracy of 4D-CT ventilation imaging: First comparison with Technegas SPECT ventilation.* Med Phys, 2017. **44**(8): p. 4045-4055.

103. Yamamoto, T., et al., *Pulmonary ventilation imaging based on 4-dimensional computed tomography: comparison with pulmonary function tests and SPECT ventilation images.* Int J Radiat Oncol Biol Phys, 2014. **90**(2): p. 414-22.

104. Bucknell, N.W., et al., *Functional lung imaging in radiation therapy for lung cancer: A systematic review and meta-analysis.* Radiother Oncol, 2018. **129**(2): p. 196-208.

105. Kamnitsas, K., et al., *Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation.* Medical Image Analysis, 2017. **36**: p. 61-78.

106. He, K.M., et al., *Deep Residual Learning for Image Recognition.* 2016 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr), 2016: p. 770-778.

107. Zhang, Y.D., et al., *Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU.* Journal of Computational Science, 2018. **28**: p. 1-10.

108. Rankine, L.J., et al., *Correlation of Regional Lung Ventilation and Gas Transfer to Red Blood Cells: Implications for Functional-Avoidance Radiation Therapy Planning.* Int J Radiat Oncol Biol Phys, 2018. **101**(5): p. 1113-1122.

109. Karimpanal, T.G. and R. Bouffanais, *Self-organizing maps for storage and transfer of knowledge in reinforcement learning.* Adaptive Behavior, 2019. **27**(2): p. 111-126.

110. Kimura, T., et al., *Combined Ventilation and Perfusion Imaging Correlates With the Dosimetric Parameters of Radiation Pneumonitis in Radiation Therapy Planning for Lung Cancer.* Int J Radiat Oncol Biol Phys, 2015. **93**(4): p. 778-87.

111. Humphries, S.M., et al., *Deep Learning Enables Automatic Classification of Emphysema Pattern at CT.* Radiology, 2020. **294**(2): p. 434-444.

112. Anthimopoulos, M., et al., *Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network.* Ieee Transactions on Medical Imaging, 2016. **35**(5): p. 1207-1216.

113. Geirhos, R., et al. *ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness.* 2018. arXiv:1811.12231.

114. Brendel, W. and M.J.a.e.-p. Bethge *Approximating CNNs with Bag-of-local-Features models works surprisingly well on ImageNet.* 2019. arXiv:1904.00760.

115. Gatys, L.A., A.S. Ecker, and M. Bethge, *Texture Synthesis Using Convolutional Neural Networks.* Advances in Neural Information Processing Systems 28 (Nips 2015), 2015. **28**.

116.   Ritter, S., et al. *Cognitive Psychology for Deep Neural Networks: A Shape Bias Case Study.* 2017. arXiv:1706.08606.

117.   LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning.* Nature, 2015. **521**(7553): p. 436-444.

118.   Kriegeskorte, N., *Deep neural networks: a new framework for modeling biological vision and brain information processing.* Perception, 2016. **45**: p. 73-73.

119.   Rieger, B., et al., *Magnetic resonance fingerprinting using echo-planar imaging: Joint quantification of T1 and T2 * relaxation times.* Magn Reson Med, 2017. **78**(5): p. 1724-1733.